

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 06:53:47 ; Search time 17 Seconds
(without alignments)
50.938 Million cell updates/sec

Title: US-10-623-176-2

Perfect score: 45

Sequence: 1 VLHDDLLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	84.4	1148	2 S51855	hypothetical prote
2	37	82.2	1165	2 D59433	C. elegans protein
3	35	77.8	119	2 T00151	hypothetical prote
4	35	77.8	604	2 S36487	E1 protein - human
5	35	77.8	605	2 S36469	E1 protein - human
6	35	77.8	616	2 C89226	type I restriction
7	34	75.6	130	2 A81316	chemotaxis regulat
8	33	73.3	145	2 A41652	probable glyceroph
9	33	73.3	174	2 C89075	protein K04A8.9 [i
10	33	73.3	210	2 T44122	probable glyceroph
11	33	73.3	210	2 E89762	hypothetical prote
12	33	73.3	238	2 S76860	hypothetical prote
13	33	73.3	249	2 A80461	glycerophosphodie
14	33	73.3	259	2 G82601	conserved hypothet
15	33	73.3	291	2 D83371	probable dehydroge
16	33	73.3	320	2 B75457	polyprenyl synthas
17	33	73.3	322	2 C83075	octaprenyl-diphosp
18	33	73.3	343	2 A10581	DNA polymerase III
19	33	73.3	343	2 B85564	DNA polymerase III
20	33	73.3	343	2 F90713	DNA polymerase III
21	33	73.3	343	2 A45251	DNA-directed DNA p
22	33	73.3	378	2 T35403	probable polypreny
23	33	73.3	397	2 S98813	hypothetical prote
24	33	73.3	398	2 H72660	probable N-Acylami
25	33	73.3	519	1 S69988	unspecific monooxy
26	33	73.3	519	1 S69988	unspecific monooxy
27	33	73.3	604	2 S36493	E1 protein - human
28	33	73.3	654	2 A80668	probable peptidase
29	33	73.3	667	2 G30883	probable collagena

30	33	73.3	667	2 H85734	probable collagena
31	33	73.3	667	2 F64895	hypothetical prote
32	33	73.3	723	2 F85047	hypothetical prote
33	33	73.3	856	2 C85023	hypothetical prote
34	33	73.3	908	2 T50695	seca protein limpo
35	33	73.3	973	2 T50449	DNA repair and rec
36	33	73.3	1127	2 S47445	MDM1 protein - yea
37	33	73.3	1770	2 S56221	hypothetical prote
38	32	71.1	91	1 C69973	ribonuclease inhib
39	32	71.1	107	1 R6KMIC	acidic ribosomal p
40	32	71.1	166	2 A42524	A-ORF-B protein.-
41	32	71.1	270	2 G84226	hypothetical prote
42	32	71.1	276	2 T28746	hypothetical prote
43	32	71.1	284	2 G70732	probable thiosulfa
44	32	71.1	296	2 T27768	hypothetical prote
45	32	71.1	297	2 A84759	probable trans-pre

ALIGNMENTS

RESULT 1

S51855
hypothetical protein YDR128w - Yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YD9302.03

C:Species: Saccharomyces cerevisiae

C:Date: 05-May-1995 #sequence_revision 21-Jul-1995 #text_change 16-Aug-2004

C:Accession: S51855

R:Oliver, K.; Harris, D.

submitted to the EMBL Data Library, February 1995

A:Reference number: S51853

A:Accession: S51855

A:Molecule type: DNA

A:Residues: 1-1148 <OLI>

A:Cross-references: UNIPROT:Q03897; EMBL:Z48179; NID:G665657; PID:G665660; GSPDB:GN000004;

C:Genetics:

A:Gene: MIPS:YDR128w

A:Cross-references: SGD:S0002535

A:Map position: 4R

C:Superfamily: WD repeat homology

Query Match 84.4%; Score 38; DB 2; Length 1148;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VLHDDLLEA 9
Db	1001	LLHDDIIEA 1009

RESULT 2

D59433
C. elegans protein Z37093 homolog [imported] - human

C:Species: Homo sapiens (man)

C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004

C:Accession: D59433; E59433

R:Nagase, T.; Seki, N.; Ishikawa, K.; Ohira, M.; Kawarabayasi, Y.; Ohara, O.; Tanaka, A.;

DNA Res. 3, 321-329, 1996

A:Title: Prediction of the coding sequences of unidentified human genes. VI. The coding s

A:Reference number: D59433; MUID:97191544; PMID:9039502

A:Accession: D59433

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1165 <NAG>

A:Cross-references: UNIPROT:Q92619; GB:BAAL3212; PID:gi504026; PIDN:BAAL3212.1

R:Ohara, O.; Nagase, T.; Kikuno, R.; Nomura, N.

submitted to Genbank, August 1996

A:Reference number: E59433

A:Accession: E59433

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1165 <STO2>

A;Cross-references: GB:BAAL3212; PID:g1504026; PIDN:BAAL3212.1

Query Match 82.2%; Score 37; DB 2; Length 1165;
Best Local Similarity 88.9%; Pred. No. 82;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
|| |||||
Db 166 VLRDDLEA 174

RESULT 3

T00151

hypothetical protein 18 - Staphylococcus aureus phage phi PVL

C;Species: Staphylococcus aureus phage phi PVL

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T00151

R;Kaneko, J.; Kimura, T.; Kawakami, Y.; Tomita, T.; Kamio, Y.

Biosci. Biotechnol. Biochem. 61, 1960-1962, 1997

A;Title: Pantone-Valentine leukocidin genes in a phage-like particle isolated from mitomy

A;Reference number: Z14119; MUID:98067870; PMID:9404084

A;Accession: T00151

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-119 <KAN>

A;Cross-references: UNIPROT:O80057; EMBL:AB009866; NID:g3341907; PIDN:BAA31891.1; PID:g3

C;Superfamily: Staphylococcus aureus phage phi PVL hypothetical protein 18

Query Match 77.8%; Score 35; DB 2; Length 119;

Best Local Similarity 85.7%; Pred. No. 15;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 HDDLLEA 9

|||:|
Db 72 HDDLEA 78

RESULT 4

S36487

E1 protein - human papillomavirus type 19

C;Species: human papillomavirus type 19

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004

C;Accession: S36487

R;Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A;Description: Primer-directed sequencing of human papillomavirus types.

A;Reference number: S36469

A;Accession: S36487

A;Molecule type: DNA

A;Residues: 1-604

A;Cross-references: UNIPROT:Q02048; EMBL:X74470; NID:g396940; PIDN:CAA52520.1; PID:g3969

C;Superfamily: papillomavirus E1 protein

C;Keywords: early protein; nucleus

Query Match 77.8%; Score 35; DB 2; Length 604;

Best Local Similarity 75.0%; Pred. No. 94;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LHDDLLEA 9

|||:|
Db 201 VHDDLLES 208

RESULT 5

S36469

E1 protein - human papillomavirus type 14D

C;Species: human papillomavirus type 14D

C;Date: 09-Dec-1993 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004

C;Accession: S36469

R;Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A;Description: Primer-directed sequencing of human papillomavirus types.

A;Reference number: S36469

A;Accession: S36469

A;Molecule type: DNA

A;Residues: 1-605

A;Cross-references: UNIPROT:P36721; EMBL:X74467; NID:g396918; PIDN:CAA52502.1; PID:g39691

A;Experimental source: strain 14D

C;Superfamily: papillomavirus E1 protein

C;Keywords: early protein; nucleus

Query Match 77.8%; Score 35; DB 2; Length 605;

Best Local Similarity 75.0%; Pred. No. 94;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LHDDLLEA 9

:|
Db 202 VHDDLLES 209

RESULT 6

C69226

type I restriction modification enzyme, subunit M - Methanobacterium thermoautotrophicum

C;Species: Methanobacterium thermoautotrophicum

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C;Accession: C69226

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F

Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.;

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi

A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: C69226

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-616 <MTH>

A;Cross-references: UNIPROT:O27025; GB:AE000868; GB:AE000666; NID:g2622025; PIDN:AAB8544

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH942

A;Start codon: GTG

C;Superfamily: type I site-specific deoxyribonuclease chain hsdM

Query Match 77.8%; Score 35; DB 2; Length 616;

Best Local Similarity 77.8%; Pred. No. 96;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9

|||:|
Db 465 VLEDDLLEA 473

RESULT 7

A81316

chemotaxis regulatory protein Cj1118c [imported] - Campylobacter jejuni (strain NCTC 1111

C;Species: Campylobacter jejuni

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 16-Aug-2004

C;Accession: A81316

R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell

Nature 403, 665-668, 2000

A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hype

A;Reference number: A81250; MUID:20150912; PMID:10688204

A;Accession: A81316

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-130 <PAR>

A;Cross-references: UNIPROT:P71129; GB:AL139077; GB:AL111168; NID:g6968444; PIDN:CAB7337;

A;Experimental source: serotype O2, strain NCTC 11168

C;Genetics:

A;Gene: cheY; Cj1118c

C;Superfamily: Signal transduction receiver (phosphoacceptor) protein, CheY type; respons

Query Match 75.6%; Score 34; DB 2; Length 130;

Best Local Similarity 85.7%; Pred. No. 26;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 06:53:02 ; Search time 81.25 Seconds
(without alignments)
56.723 Million cell updates/sec

Title: US-10-623-176-2

Perfect score: 45

Sequence: 1 VLHDDLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	100.0	1136	2	Q6P189	Q6P189 homo sapien
2	40	88.9	571	2	Q6KZHS	Q6KZHS picophilus
3	39	86.7	668	2	Q6LZ81	Q6LZ81 methanococc
4	38	84.4	1148	2	Q03897	Q03897 saccharomyc
5	37	82.2	1131	2	Q81YN3	Q81YN3 homo sapien
6	37	82.2	1136	2	Q8HX84	Q8HX84 homo sapien
7	37	82.2	1165	2	Q22619	Q22619 homo sapien
8	36	80.0	101	2	Q6AMV9	Q6AMV9 desulfotale
9	36	80.0	231	2	Q6AUN7	Q6AUN7 oryza sativ
10	36	80.0	585	2	Q8EJS9	Q8EJS9 shewanella
11	36	80.0	691	2	Q8PQRO	Q8PQRO xanthomonas
12	35	77.8	119	2	Q80057	Q80057 staphylococ
13	35	77.8	120	2	Q8SCW6	Q8SCW6 pseudomonas
14	35	77.8	132	2	Q7MSU3	Q7MSU3 mouse adeno
15	35	77.8	148	2	Q6ZW54	Q6ZW54 homo sapien
16	35	77.8	376	2	Q8MVB1	Q8MVB1 ixodes scap
17	35	77.8	496	2	Q8SDK3	Q8SDK3 staphylococ
18	35	77.8	496	2	Q9MBP0	Q9MBP0 staphylococ
19	35	77.8	603	1	V81_HPV21	V81_HPV21 human papil
20	35	77.8	604	1	V81_HPV19	V81_HPV19 human papil
21	35	77.8	605	1	V81_HPV14	V81_HPV14 human papil
22	35	77.8	616	2	Q8JH28	Q8JH28 methanobact
23	35	77.8	843	2	Q8JH28	Q8JH28 gallus gall
24	35	77.8	872	2	Q6DIR8	Q6DIR8 xenopus tro
25	34	75.6	68	2	Q8EYR5	Q8EYR5 leptospira
26	34	75.6	130	1	CHEY_CAMJTB	CHEY_CAMJTB campylobact
27	34	75.6	146	2	Q74B30	Q74B30 geobacter s
28	34	75.6	195	2	Q7PTL2	Q7PTL2 anopheles g
29	34	75.6	228	2	Q8RJW9	Q8RJW9 streptomyce
30	34	75.6	241	2	Q6DVA9	Q6DVA9 gibberella
31	34	75.6	248	2	Q6DV84	Q6DV84 bacillus sp

32	34	75.6	253	2	Q889T7	Q889T7 pseudomonas
33	34	75.6	262	2	Q84GC9	Q84GC9 uncultured
34	34	75.6	268	2	Q9BJL6	Q9BJL6 giardia lam
35	34	75.6	293	2	Q7QYU3	Q7QYU3 giardia lam
36	34	75.6	366	2	Q65CP2	Q65CP2 bacillus li
37	34	75.6	372	2	Q87ZG8	Q87ZG8 pseudomonas
38	34	75.6	401	2	Q6K4V7	Q6K4V7 oryza sativ
39	34	75.6	408	2	Q94LF5	Q94LF5 oryza sativ
40	34	75.6	408	2	Q84MW4	Q84MW4 oryza sativ
41	34	75.6	471	2	Q8NJ19	Q8NJ19 mucor circi
42	34	75.6	489	2	Q8MS72	Q8MS72 drosophila
43	34	75.6	529	2	Q7PRG2	Q7PRG2 anopheles g
44	34	75.6	536	2	Q7Q5X6	Q7Q5X6 giardia lam
45	34	75.6	640	2	Q9NKM2	Q9NKM2 leishmania

ALIGNMENTS

RESULT 1

Q6P189 PRELIMINARY; PRT; 1136 AA.
AC Q6P189;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
GN Minor histocompatibility antigen HA-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
EX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065223; AAH65223.1; -.
DR HSSP; Q07960; 1AM4.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001060; Cdc15 Fes Cip4.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR002219; DAG_FE-bind.
DR InterPro; IPR000198; RhoGAP.
DR InterPro; IPR008936; Rho_GAP.
DR Pfam; PF00130; Cl 1; 1.
DR Pfam; PF00620; RhoGAP; 1.
DR SMART; SM00109; Cl; 1.
DR SMART; SM00055; FCH; 1.

```

DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS00479; DAG PE BIND DOM 1; UNKNOWN_1.
DR PROSITE; PS50081; DAG PE_BIND_DOM_2; 1.
DR PROSITE; PS50238; RHO GAP; 1.
SQ SEQUENCE 1136 AA; 134550 MW; 14B7073296102DD5 CRC64;

Query Match      100.0%; Score 45; DB 2; Length 1136;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
Db 137 VLHDDLLEA 145

RESULT 2
Q6KZH5 PRELIMINARY; PRT; 571 AA.
AC Q6KZH5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Molybdopterin biosynthesis MoeA protein.
GN OrderedLocustNames=PT01292;
OS Picrophilus torridus.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Picrophilaceae; Picrophilus.
OX NCBI_TaxID=82076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 9790 / ATCC 700027;
RX PubMed=15184674; DOI=10.1073/pnas.0401356101;
RA Fuetterer O., Angelov A., Liesegang H., Gertschalk G., Schleper C.,
RA Schepers B., Dock C., Antranikian G., Liebl W.;
RT "Genome sequence of Picrophilus torridus and its implications for life
RT around pH 0.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9091-9096(2004).
RL EMBL; AE017261; AAT43877.1; -.
GO: GO:0006777; P-Mo-molybdopterin cofactor biosynthesis; IEA.
DR InterPro; IPR001453; MoCF_biosynth.
DR InterPro; IPR005111; MoeA_C.
DR Pfam; PF00994; MoCF_biosynth; 1.
DR Pfam; PF03454; MoeA_C; 1.
DR Pfam; PF03453; MoeA_N; 1.
DR ProDom; PD002460; MoCF_biosynth; 1.
DR TIGRFAMs; TIGR00177; molyb_syn; 1.
KW Complete proteome.
SQ SEQUENCE 571 AA; 64707 MW; D2EB18FBFC115FA1 CRC64;

Query Match      88.9%; Score 40; DB 2; Length 571;
Best Local Similarity 88.9%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
Db 63 VLHDDLLEA 71

RESULT 3
Q6LZ81 PRELIMINARY; PRT; 668 AA.
AC Q6LZ81;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MCM family related protein.
GN OrderedLocustNames=MP0748;
OS Methanococcus maripaludis.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanococcaceae; Methanococcus.
OX NCBI_TaxID=39152;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2 / LL;
RA Hendrickson E.L., Kaul R., Zhou Y., Bovee D., Chapman P., Chung J.,
RA Conway de Macario E., Dodsworth J., Gillett W., Graham D.E.,
RA Haydock A.K., Kang A., Land M.L., Levy R., Lie T.J., Major T.,
RA Moore B., Porat I., Overbeek R., Palmeiri A., Rouse G.,
RA Saenphimmachak C., Soll D., Whitman W.B., Larimer F.W., Olson M.V.,
RA Leigh J.A.;
RT "Complete genome sequence of the mesophilic hydrogenotrophic
RT methanogen Methanococcus maripaludis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX957221; CAP30304.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO: GO:0006270; F:DNA replication initiation; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001208; MCM.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR PRINTS; PR01657; MCMFAMILY.
DR ProDom; PD001041; MCM; 1.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS00051; MCM_2; 1.
KW Complete proteome.
SQ SEQUENCE 668 AA; 75660 MW; 2ADAF3800B1049F4 CRC64;

Query Match      86.7%; Score 39; DB 2; Length 668;
Best Local Similarity 88.9%; Pred. No. 82;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
Db 628 VLHDDLLEA 636

RESULT 4
Q03897 PRELIMINARY; PRT; 1148 AA.
AC Q03897;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN ORFNames=YDR128W;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB972;
RA Oliver K., Harris D.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB972;
RA Barrell B., Rajandream M.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 4 WD repeats.
DR EMBL; Z48179; CAA88209.1; -.
DR PIR; S51855; S51855.
DR GO: GO:0000253; YDR128W.
DR GO: GO:0000329; C:vacuolar membrane (sensu Fungi); IDA.
DR InterPro; IPR006575; RWD.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 4.
DR PRINTS; PR00320; GPROTEINRPT.
DR SMART; SM00591; RWD; 1.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PSS0098; RWD; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 06:44:47 ; Search time 90.5 Seconds
(without alignments)
38.462 Million cell updates/sec

Title: US-10-623-176-2

Perfect score: 45

Sequence: 1 VLHDDLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_l6Dec04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	45	100.0	9	2	AAW99196	Aaw99196 Minor his
2	45	100.0	9	2	AAW97375	Aaw97375 HA-1 H-al
3	45	100.0	9	8	ADH40333	Adh40333 Human min
4	45	100.0	13	2	AAW99199	Aaw99199 vR cell K
5	45	100.0	13	2	AAW97415	Aaw97415 KIAA0223
6	45	100.0	1165	8	ADN11768	Adn11768 Human the
7	38	84.4	199	7	ADF06830	Adf06830 Bacterial
8	37	82.2	9	2	AAW99197	Aaw99197 Minor his
9	37	82.2	9	2	AAW97374	Aaw97374 HA-1 R-al
10	37	82.2	9	8	ADH40334	Adh40334 Human min
11	37	82.2	13	2	AAW99198	Aaw99198 DH cell K
12	37	82.2	13	2	AAW97414	Aaw97414 KIAA0223
13	37	82.2	1136	5	ADI11793	Adi11793 Human NOV
14	37	82.2	1136	5	ADI16744	Adi16744 Human NOV
15	37	82.2	1136	8	ADH61299	Adh61299 INTSIG pr
16	37	82.2	1136	8	ADN42398	Adn42398 Human nov
17	37	82.2	1136	8	ADS10490	Ads10490 Human the
18	37	82.2	1165	5	ADI117194	Adi117194 Human NOV
19	37	82.2	1165	8	ADP55035	Adp55035 Human PRO
20	37	82.2	1165	8	ADP24600	Adp24600 PRO polyP
21	36	80.0	9	2	AAW99195	Aaw99195 Minor his
22	36	80.0	9	2	AAW97572	Aaw97572 T-cell ep
23	35	77.8	148	8	ADQ65728	Adq65728 Novel hum
24	35	77.8	496	6	ABM71411	Abm71411 Staphyloc
25	35	77.8	608	7	ABO77464	AbO77464 Pseudomon

26	35	77.8	785	6	ABU29875	Abu29875 Protein e
27	35	77.8	789	7	ADC97107	Adc97107 E. faeciu
28	35	77.8	964	7	ADB70281	Adb70281 C. neoFor
29	34	75.6	130	6	ABU26589	Abu26589 Protein e
30	34	75.6	232	4	AAU55593	Aau55593 Propionib
31	34	75.6	232	6	ABM52112	Abm52112 Propionib
32	34	75.6	253	6	ABU42064	Abu42064 Protein e
33	34	75.6	332	4	ABB64157	Abb64157 Drosophil
34	34	75.6	385	6	ABU41832	Abu41832 Protein e
35	34	75.6	440	4	AAAB74623	AaB74623 Saitoella
36	34	75.6	446	6	ABU34950	Abu34950 Protein e
37	34	75.6	680	8	ADJ48425	Adj48425 Maize oil
38	34	75.6	774	4	ABB65640	Abb65640 Drosophil
39	34	75.6	774	4	AAU38963	Aau38963 Drosophil
40	34	75.6	774	7	ADC35871	Adc35871 Drosophil
41	34	75.6	783	4	AAG91263	Aag91263 C glutami
42	34	75.6	802	7	AAE38199	Aae38199 Fruit fly
43	34	75.6	1383	4	ABB63586	Abb63586 Drosophil
44	33	73.3	70	7	ADH85816	Adh85816 Enterococ
45	33	73.3	124	6	ABU31332	Abu31332 Protein e

ALIGNMENTS

RESULT 1		
AAW99196		
ID	AAW99196	standard; peptide; 9 AA.
XX	AC	AAW99196;
XX	DT	20-MAY-1999 (first entry)
XX	DE	Minor histocompatibility antigen HA-1 T-cell epitope #2.
XX	KW	Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
KW	KW	graft versus host disease; bone marrow transplant; leukaemia; vaccine;
KW	KW	diagnosis; aplastic anaemia; immune deficiency disease.
XX	XX	
XX	XX	
OS	Homo sapiens.	
XX	.	
XX	WO9905174-A1.	
PN	XX	
XX	XX	
PD	04-FEB-1999.	
XX	XX	
XX	23-JUL-1998;	98WO-NL000425.
PF		
XX	23-JUL-1997;	97EP-00202303.
XX		
PR		

Query Match 100.0%; Score 45; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDDLLEA 9
DB 1 VLHDDLLEA 9

RESULT 2
AAW97375
ID AAW97375 standard; protein; 9 AA.
XX
AC AAW97375;
XX
DT 13-MAY-1999 (first entry)
XX
XX HA-1 H-allele sequence.
XX
KW Intron; minor histocompatibility antigen HA-1; typing allele; H allele;
KW R allele; polymorphic nucleotide; HA-1 typing; bone marrow transplant;
KW severe aplastic anaemia; leukaemia; immune deficiency disease; ss.
XX
OS Homo sapiens.
XX
PN WO9905313-A2.
XX
PD 04-FEB-1999.
XX
PF 23-JUL-1998; 98WO-EP004928.
XX
PR 23-JUL-1997; 97EP-00202303.
PR 02-JUN-1998; 98EP-00870125.
XX
PA (UYLE-) RIJKSUNIV LEIDEN.
XX
PI Goulmy E;
XX
DR WPI; 1999-142960/12.
XX
XX Typing minor histocompatibility antigen HA-1 - by amplifying and
PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g. detection
PT of genetic aberrances.
XX
PS Claim 18; Fig 5; 59pp; English.
XX
CC The present sequence represents part of the minor histocompatibility
CC antigen HA-1 H-allele. The specification describes methods for typing
CC alleles (preferably the H and R alleles) of the minor histocompatibility
CC antigen HA-1 in a sample, which comprise detecting polymorphic
CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The
CC methods can be used for HA-1 typing for bone marrow transplants, severe
CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as
CC detection of genetic aberrances. The probes and primers of the invention
CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be used
CC anti-idiotypic B cells and/or T cells and antibodies
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDDLLEA 9
DB 1 VLHDDLLEA 9

RESULT 3
ADH40333
ID ADH40333 standard; peptide; 9 AA.
XX
AC ADH40333;
XX

DT 11-MAR-2004 (first entry)
XX
DE Human minor histocompatibility antigen HA-1 T cell epitope.
XX
XX human; cytostatic; vaccine; SNP profile; cancer; leukaemia;
KW minor histocompatibility antigen; mHAG; T cell epitope.
XX
OS Homo sapiens.
XX
PN WO2003106692-A2.
XX
PD 24-DEC-2003.
XX
PF 13-JUN-2003; 2003WO-EP006251.
XX
PR 13-JUN-2002; 2002EP-00013423.
XX
PA (MERE) MERCK PATENT GMBH.
XX
PI Strittmatter W, Moll H;
XX
DR WPI; 2004-082200/08.
XX
XX Providing allelic variant epitope of protein based on single nucleotide
PT polymorphism by defining target protein, screening database of protein,
PT identifying, selecting allelic variant protein, creating variant
PT epitopes.
XX
PS Disclosure; Page 82; 119pp; English.
XX
CC The invention relates to a novel method for providing epitopes of allelic
CC variants of antigenic proteins from specific species based on single
CC nucleotide polymorphism (SNP), by defining target protein/peptide or its
CC subset, screening database of DNA encoding target protein, identifying,
CC selecting allelic peptide/protein variants, expression product or its
CC fragment encoded by DNA sequence having SNP, creating variant epitopes,
CC selecting epitopes binding to MHC protein. A protein of the invention has
CC cytostatic activity, and may have a use in a vaccine. The method is
CC useful for generating a SNP profile of one or more individuals from a
CC given species by applying the method for several protein from the
CC individuals, where the SNP profile was related to disease, preferably
CC generating the SNP-related polymorphic profile. A method of the invention
CC is useful for transplanting haematopoietic stem cells from a donor to a
CC recipient and treating cancer, preferably leukaemia, and for determining
CC the progression, regression or onset of a treated disease. The present
CC sequence is used in the exemplification of the invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDDLLEA 9
DB 1 VLHDDLLEA 9

RESULT 4
AAW99199
ID AAW99199 standard; peptide; 13 AA.
XX
AC AAW99199;
XX
DT 20-MAY-1999 (first entry)
XX
DE vR cell KIAA0223 protein sequence.
XX
KW Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;
KW diagnosis; aplastic anaemia; immune deficiency disease.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:09:04 ; Search time 70 Seconds
(without alignments)
44.445 Million cell updates/sec

Title: US-10-623-176-2

Perfect score: 45

Sequence: 1 VLHDDLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 1465611

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pap:*
- 2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pap:*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pap:*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pap:*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pap:*
- 6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pap:*
- 7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pap:*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pap:*
- 9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pap:*
- 10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pap:*
- 11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pap:*
- 12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pap:*
- 13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pap:*
- 14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pap:*
- 15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pap:*
- 16: /cgn2_6/ptodata/1/pubaa/US10D_PUBCOMB.pap:*
- 17: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pap:*
- 18: /cgn2_6/ptodata/1/pubaa/US11_NEW_PUB.pap:*
- 19: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pap:*
- 20: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	9	15	US-10-623-176-2
2	45	100.0	9	16	US-10-791-217-2
3	45	100.0	9	17	US-10-861-335-1
4	45	100.0	10	15	US-10-623-176-43
5	45	100.0	13	15	US-10-623-176-76
6	45	100.0	13	16	US-10-791-217-16
7	45	100.0	17	15	US-10-623-176-57
8	45	100.0	22	15	US-10-623-176-53
9	45	100.0	25	15	US-10-623-176-55
10	45	100.0	27	15	US-10-623-176-51
11	45	100.0	29	15	US-10-623-176-49
12	41	91.1	9	15	US-10-623-176-41
13	41	91.1	9	15	US-10-623-176-45

14	39	86.7	65	15	US-10-424-599-242306	Sequence 242306,
15	37	82.2	9	15	US-10-623-176-10	Sequence 10, Appl
16	37	82.2	9	15	US-10-623-176-47	Sequence 47, Appl
17	37	82.2	9	16	US-10-791-217-5	Sequence 5, Appl
18	37	82.2	10	15	US-10-623-176-44	Sequence 44, Appl
19	37	82.2	12	15	US-10-623-176-65	Sequence 65, Appl
20	37	82.2	13	15	US-10-623-176-74	Sequence 74, Appl
21	37	82.2	13	16	US-10-791-217-14	Sequence 14, Appl
22	37	82.2	17	15	US-10-623-176-69	Sequence 69, Appl
23	37	82.2	19	15	US-10-623-176-52	Sequence 52, Appl
24	37	82.2	19	15	US-10-623-176-70	Sequence 70, Appl
25	37	82.2	21	15	US-10-623-176-71	Sequence 71, Appl
26	37	82.2	22	15	US-10-623-176-63	Sequence 63, Appl
27	37	82.2	23	15	US-10-623-176-72	Sequence 72, Appl
28	37	82.2	25	15	US-10-623-176-67	Sequence 67, Appl
29	37	82.2	26	15	US-10-623-176-66	Sequence 66, Appl
30	37	82.2	27	15	US-10-623-176-60	Sequence 60, Appl
31	37	82.2	28	15	US-10-623-176-61	Sequence 61, Appl
32	37	82.2	29	15	US-10-623-176-59	Sequence 59, Appl
33	37	82.2	1136	15	US-10-072-012-280	Sequence 280, App
34	37	82.2	1136	15	US-10-072-012-729	Sequence 729, App
35	37	82.2	1165	15	US-10-072-012-730	Sequence 730, App
36	36	80.0	9	15	US-10-623-176-1	Sequence 1, Appl
37	36	80.0	9	15	US-10-623-176-40	Sequence 40, Appl
38	36	80.0	9	16	US-10-791-217-1	Sequence 1, Appl
39	36	80.0	10	15	US-10-623-176-6	Sequence 6, Appl
40	36	80.0	10	15	US-10-623-176-24	Sequence 24, Appl
41	36	80.0	13	15	US-10-623-176-54	Sequence 54, Appl
42	36	80.0	17	15	US-10-623-176-33	Sequence 33, Appl
43	36	80.0	50	15	US-10-424-599-198061	Sequence 198061,
44	36	80.0	231	16	US-10-437-963-204496	Sequence 204496,
45	36	80.0	922	16	US-10-437-963-137943	Sequence 137943,

ALIGNMENTS

RESULT 1
US-10-623-176-2
; Sequence 2, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Els A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-604705
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(9)
US-10-623-176-2

Query Match 100.0%; Score 45; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
| | | | |
Db 1 VLHDDLLEA 9

RESULT 2

US-10-791-217-2
; Sequence 2, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-6479US
; CURRENT APPLICATION NUMBER: US/10/791,217
; PRIOR FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: obtained from histocompatibility antigen
US-10-791-217-2

Query Match 100.0%; Score 45; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
| | | | |
Db 1 VLHDDLLEA 9

RESULT 3

US-10-861-335-1
; Sequence 1, Application US/10861335
; Publication No. US20050031612A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; TITLE OF INVENTION: Minor histocompatibility antigen HA-1: target antigen for immunot
; FILE REFERENCE: 2183-6479US
; CURRENT APPLICATION NUMBER: US/10/861,335
; CURRENT FILING DATE: 2004-06-04
; PRIOR APPLICATION NUMBER: PCT/NL02/00791
; PRIOR FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: EP 01204704.9
; PRIOR FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HA-1 peptide
US-10-861-335-1

Query Match 100.0%; Score 45; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
| | | | |
Db 1 VLHDDLLEA 9

RESULT 4

US-10-623-176-43
; Sequence 43, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Els A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 43
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; NAME/KEY: SITE
; LOCATION: (1)..(10)
US-10-623-176-43

Query Match 100.0%; Score 45; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
| | | | |
Db 1 VLHDDLLEA 9

RESULT 5

US-10-623-176-76
; Sequence 76, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Els A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 76
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: KIAA0223
; OTHER INFORMATION: sequence derived from a presumed HA-1 homozygous
US-10-623-176-76

Query Match 100.0%; Score 45; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 06:54:48 ; Search time 24 Seconds
(without alignments)
27.993 Million cell updates/sec

Title: US-10-623-176-2

Perfect score: 45

Sequence: 1 VLHDDLLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCITUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	9	4	US-09-269-250E-20
2	45	100.0	13	4	US-09-269-250E-28
3	45	100.0	26	4	US-09-269-250E-36
4	38	84.4	199	4	US-09-543-681A-7115
5	38	84.4	1148	4	US-09-538-092-156
6	37	82.2	9	4	US-09-269-250E-18
7	37	82.2	13	4	US-09-269-250E-26
8	36	80.0	9	4	US-09-269-250E-29
9	35	77.8	608	4	US-09-252-991A-26210
10	35	77.8	789	4	US-09-107-532A-6734
11	34	75.6	435	4	US-09-270-767-33480
12	34	75.6	435	4	US-09-270-767-48697
13	34	75.6	446	4	US-09-830-111E-2
14	33	73.3	70	4	US-09-134-000C-3701
15	33	73.3	256	3	US-09-134-001C-4544
16	33	73.3	295	4	US-09-252-991A-24727
17	33	73.3	332	4	US-09-902-540-12129
18	33	73.3	341	4	US-09-252-991A-21448
19	33	73.3	342	4	US-09-818-780-97
20	33	73.3	343	1	US-08-279-058B-9
21	33	73.3	343	4	US-08-828-323-9
22	33	73.3	352	4	US-09-489-039A-12951
23	33	73.3	623	4	US-09-710-279-1068
24	33	73.3	687	4	US-09-489-039A-9868
25	33	73.3	851	4	US-09-071-035-326
26	33	73.3	851	4	US-09-071-035-330
27	33	73.3	851	4	US-09-071-035-334

28	33	73.3	1554	4	US-09-252-991A-26814	Sequence 26814, A
29	33	73.3	4150	3	US-09-428-517-2	Sequence 2, Appli
30	32	71.1	35	3	US-09-082-279B-454	Sequence 454, App
31	32	71.1	35	3	US-09-082-279B-455	Sequence 455, App
32	32	71.1	35	3	US-09-082-279B-456	Sequence 456, App
33	32	71.1	35	3	US-08-474-349A-481	Sequence 481, App
34	32	71.1	35	3	US-08-474-349A-482	Sequence 482, App
35	32	71.1	35	3	US-08-474-349A-483	Sequence 483, App
36	32	71.1	35	3	US-09-315-304B-454	Sequence 454, App
37	32	71.1	35	3	US-09-315-304B-455	Sequence 455, App
38	32	71.1	35	3	US-09-315-304B-456	Sequence 456, App
39	32	71.1	35	4	US-09-834-784-454	Sequence 454, App
40	32	71.1	35	4	US-09-834-784-455	Sequence 455, App
41	32	71.1	35	4	US-09-834-784-456	Sequence 456, App
42	32	71.1	35	4	US-09-515-965A-454	Sequence 454, App
43	32	71.1	35	4	US-09-515-965A-455	Sequence 455, App
44	32	71.1	35	4	US-09-515-965A-456	Sequence 456, App
45	32	71.1	35	4	US-09-350-641C-454	Sequence 454, App

ALIGNMENTS

RESULT 1

US-09-269-250E-20
; Sequence 20, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Exon Fragments
US-09-269-250E-20

Query Match 100.0%; Score 45; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
| | | | | | | | | |
Db 1 VLHDDLLEA 9

RESULT 2

US-09-269-250E-28
; Sequence 28, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PCR Product
US-09-269-250E-28

Query Match 100.0%; Score 45; DB 4; Length 13;

```
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurPatSeqFormatter Version 0.9
; SEQ ID NO 156
; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number YDR128W
US-09-538-092-156

Query Match      84.4%; Score 38; DB 4; Length 1148;
Best Local Similarity 66.7%; Pred. No. 75;
Matches      6; Conservative      3; Mismatches      0; Indels      0; Gaps      0;

Qy      1 VLHDDLLEA 9
Db      1001 VLHDDILEA 1009

RESULT 6
US-09-269-250E-18
; Sequence 18, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Exon Fragments
US-09-269-250E-18

Query Match      82.2%; Score 37; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches      8; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

Qy      1 VLHDDLLEA 9
Db      1 VLHDDLLEA 9

RESULT 7
US-09-269-250E-26
; Sequence 26, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 13
; TYPE: PRT
```

```
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurPatSeqFormatter Version 0.9
; SEQ ID NO 156
; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number YDR128W
US-09-538-092-156

Query Match      84.4%; Score 38; DB 4; Length 1148;
Best Local Similarity 66.7%; Pred. No. 75;
Matches      6; Conservative      3; Mismatches      0; Indels      0; Gaps      0;

Qy      1 VLHDDLLEA 9
Db      1001 VLHDDILEA 1009

RESULT 6
US-09-269-250E-18
; Sequence 18, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PCR Product
US-09-269-250E-36

Query Match      100.0%; Score 45; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      1 VLHDDLLEA 9
Db      16 VLHDDLLEA 24

RESULT 4
US-09-543-681A-7115
; Sequence 7115, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7115
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7115

Query Match      84.4%; Score 38; DB 4; Length 199;
Best Local Similarity 77.8%; Pred. No. 11;
Matches      7; Conservative      1; Mismatches      1; Indels      0; Gaps      0;

Qy      1 VLHDDLLEA 9
Db      178 VLHDDLIRA 186

RESULT 5
US-09-538-092-156
; Sequence 156, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 06:53:47 ; Search time 17 Seconds
(without alignments)
50.938 Million cell updates/sec

Title: US-10-623-176-10

Perfect score: 42

Sequence: 1 VLRDDLLLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	1165	2 D59433	C. elegans protein
2	36	85.7	174	2 C89075	protein K04A8.9 [i
3	35	83.3	616	2 C69226	type I restriction
4	34	81.0	192	2 A12271	transcription regu
5	34	81.0	282	2 S31258	probable membrane
6	34	81.0	343	2 A64443	glyceroldehyde-3-p
7	34	81.0	389	2 T20604	hypothetical prote
8	34	81.0	486	2 A82878	replicative DNA he
9	34	81.0	3519	2 S43048	polyketide synthas
10	33	78.6	287	2 AG3118	ABC transporter, m
11	33	78.6	288	2 H38168	hypothetical prote
12	33	78.6	300	2 D83714	surface adhesin A
13	33	78.6	324	2 T11672	quinone oxidoreduc
14	33	78.6	483	2 H85073	probable transposo
15	33	78.6	822	2 I38728	epidermal growth f
16	33	78.6	1041	2 E70760	probable iLES prot
17	33	78.6	1059	2 E87058	isoleucyl-tRNA syn
18	32	76.2	237	2 D70032	two-component resp
19	32	76.2	264	2 T41578	hypothetical prote
20	32	76.2	335	2 AC0786	nucleoid-associated
21	32	76.2	335	2 A64988	hypothetical 37.8
22	32	76.2	335	2 H85857	nucleoid-associate
23	32	76.2	335	2 F91013	nucleoid-associate
24	32	76.2	372	2 D70753	probable oxidoredu
25	32	76.2	389	2 T34766	iron-sulfur cofact
26	32	76.2	432	2 AF2801	hypothetical prote
27	32	76.2	432	2 H97580	hypothetical prote
28	32	76.2	452	2 T40769	hypothetical prote
29	32	76.2	481	2 AD1337	transcription regu

30	32	76.2	499	2 GB6538	leucyl aminopeptid
31	32	76.2	499	2 G72083	leucine aminopepti
32	32	76.2	525	2 AF2950	GGDEF family prote
33	32	76.2	525	2 E98332	hypothetical prote
34	32	76.2	545	2 T00485	probable phosphori
35	32	76.2	568	2 DB3182	hypothetical prote
36	32	76.2	587	2 E71020	hypothetical prote
37	32	76.2	599	1 A34231	sulfite reductase
38	32	76.2	599	1 H65057	sulfite reductase
39	32	76.2	599	2 AG0858	sulfite reductase
40	32	76.2	599	2 C91081	sulfite reductase
41	32	76.2	599	2 DB5926	sulfite reductase
42	32	76.2	606	2 AF0409	sulfite reductase
43	32	76.2	826	2 T46080	hypothetical prote
44	32	76.2	826	2 T46061	hypothetical prote
45	32	76.2	889	2 T02240	nitrate reductase

ALIGNMENTS

RESULT 1

D59433
C: elegans protein Z37093 homolog [imported] - human
C:Species: Homo sapiens (man)
C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C:Accession: D59433; E59433
R:Nagase, T.; Seki, N.; Ishikawa, K.; Ohira, M.; Kawarabayasi, Y.; Ohara, O.; Tanaka, A.;
DNA Res. 3, 321-329, 1996
A>Title: Prediction of the coding sequences of unidentified human genes. VI. The coding s

A:Reference number: D59433; MUID:97191544; PMID:9039502

A:Accession: D59433

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1165 <NAG>

A:Cross-references: UNIPROT:Q92619; GB:BAAL3212; PID:g1504026; PIDN:BAAL3212.1

R:Ohara, O.; Nagase, T.; Kikuno, R.; Nomura, N.

submitted to Genbank, August 1996

A:Reference number: E59433

A:Accession: E59433

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1165 <STO2>

A:Cross-references: GB:BAAL3212; PID:g1504026; PIDN:BAAL3212.1

Query Match 100.0%; Score 42; DB 2; Length 1165;

Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRDDLLLEA 9

Db 166 VLRDDLLLEA 174

RESULT 2

C89075
protein K04A8.9 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C89075
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_eleg
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: C89075
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-174 <STO>
A:Cross-references: UNIPROT:Q94256; GB:chr_V; PIDN:NAAC48055.1; PID:g1658358; GSPDB:GN000;

A;Gene: K04A8.9
A;Map position: 5

Query Match 85.7%; Score 36; DB 2; Length 174;
Best Local Similarity 77.8%; Pred. No. 9.9;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLKDDLLLEA 9
||:||||:|
Db 129 VLKDDLLLEA 137

RESULT 3

C69226
Type I restriction modification enzyme, subunit M - Methanobacterium thermoautotrophicum
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: C69226
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Mierzbowski, J.; Gibson, R.; Jiواني, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: C69226
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-616 <MTH>
A;Cross-references: UNIPROT:O27025; GB:AS000868; GB:AS000666; NID:G2622025; PIDN:AAB8544
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH942
A;Start codon: GTG
C;Superfamily: type I site-specific deoxyribonuclease chain hsdM

Query Match 83.3%; Score 35; DB 2; Length 616;
Best Local Similarity 77.8%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLKDDLLLEA 9
||:||||:|
Db 465 VLEDDLLLEA 473

RESULT 4

AI2271
transcription regulator all3728 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AI2271
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AI2271
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-192 <KUR>
A;Cross-references: UNIPROT:Q8YQT9; GB:BA000019; PIDN:BA875427.1; PID:G17132862; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all3728

Query Match 81.0%; Score 34; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RDDLLLEA 9
|||||||
Db 60 RDDLLLEA 66

RESULT 5

S31258
probable membrane protein FUN34 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein N2029; protein YNR002C
C;Species: Saccharomyces cerevisiae
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: S31258; S45126; S48340; S63328
R;Stettler, S.; Mariotte, S.; Riva, M.; Sentenac, A.; Thuriaux, P.
J. Biol. Chem. 267, 21390-21395, 1992
A;Title: An essential and specific subunit of RNA polymerase III (C) is encoded by gene F
A;Reference number: A45107; MUID:93016077; PMID:1400451
A;Accession: S31258
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-282 <STE>
A;Cross-references: UNIPROT:P32907; EMBL:X63746; NID:G4378; PIDN:CAA45279.1; PID:G4380
R;Verhasselt, P.; Aert, R.; Voet, M.; Volckaert, G.
submitted to the EMBL Data Library, January 1994
A;Description: Twelve open reading frames revealed on the 23.6 kbp segment flanking the c
A;Reference number: S45119
A;Accession: S45126
A;Molecule type: DNA
A;Residues: 1-282 <VER>
A;Cross-references: EMBL:X77395; NID:G496717; PIDN:CAA54571.1; PID:G496720
R;Verhasselt, P.; Aert, R.; Voet, M.; Volckaert, G.
Yeast 10, 1355-1361, 1994
A;Title: Twelve open reading frames revealed in the 23.6 kb segment flanking the centrom
A;Reference number: S48338; MUID:95208356; PMID:7900425
A;Accession: S48340
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-282 <VE2>
A;Cross-references: EMBL:X77395; NID:G496717; PIDN:CAA54571.1; PID:G496720
R;Aert, R.; Verhasselt, P.; Voet, M.; Volckaert, G.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S62910
A;Accession: S63328
A;Molecule type: DNA
A;Residues: 1-282 <ABR>
A;Cross-references: EMBL:Z71617; NID:G1302470; PIDN:CAA96278.1; PID:G1302471; MIPS:YNR00;
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:FUN34
A;Cross-references: SGD:S0005285; MIPS:YNR002C
A;Map position: 14R
C;Superfamily: Saccharomyces cerevisiae probable membrane protein FUN34
C;Keywords: transmembrane protein
F;119-135/Domain: transmembrane #status predicted <TM1>
F;186-202/Domain: transmembrane #status predicted <TM2>
F;209-225/Domain: transmembrane #status predicted <TM3>
F;234-225/Domain: transmembrane #status predicted <TM4>

Query Match 81.0%; Score 34; DB 2; Length 282;
Best Local Similarity 87.5%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LRDDLLLEA 9
|||||||
Db 62 LRDDLLLEA 69

RESULT 6

A64443
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Methanococcus
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: A64443
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, S.
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurt, M.A.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 06:53:02 ; Search time 81.25 Seconds
(without alignments)
56.723 Million cell updates/sec

Title: US-10-623-176-10
Perfect score: 42
Sequence: 1 VLRDDLEA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues 1612378
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	42	100.0	1131	2	Q8IYN3	Q8IYN3 homo sapien
2	42	100.0	1136	2	Q8HX84	Q8HX84 homo sapien
3	42	100.0	1165	2	Q92619	Q92619 homo sapien
4	40	95.2	376	2	Q8MVB1	Q8MVB1 ixodes scap
5	38	90.5	4150	2	Q9KIV4	Q9KIV4 streptomyce
6	37	88.1	1136	2	Q8P189	Q8P189 homo sapien
7	37	88.1	1186	2	Q8JH78	Q8JH78 brachydanio
8	36	85.7	174	2	Q94256	Q94256 caenorhabdi
9	36	85.7	585	2	Q8EJ59	Q8EJ59 shewanella
10	35	83.3	223	2	Q57K14	Q57K14 symbiobacte
11	35	83.3	361	2	Q6ALW3	Q6ALW3 desulfotale
12	35	83.3	381	2	Q75ND9	Q75ND9 lactarius c
13	35	83.3	616	2	Q270Z5	Q270Z5 methanobact
14	35	83.3	668	2	Q6LZ81	Q6LZ81 methanococc
15	35	83.3	696	2	Q65TH9	Q65TH9 mannaemia
16	34	81.0	192	2	Q8YQ79	Q8YQ79 anabaena sp
17	34	81.0	209	2	Q8B0A5	Q8B0A5 influenza a
18	34	81.0	211	2	Q8B0B1	Q8B0B1 influenza a
19	34	81.0	211	2	Q8B0B3	Q8B0B3 influenza a
20	34	81.0	212	2	Q8B099	Q8B099 influenza a
21	34	81.0	212	2	Q8B0A7	Q8B0A7 influenza a
22	34	81.0	212	2	Q8B0A9	Q8B0A9 influenza a
23	34	81.0	213	2	Q8B0A1	Q8B0A1 influenza a
24	34	81.0	213	2	Q8B0B5	Q8B0B5 influenza a
25	34	81.0	213	2	Q8B0B9	Q8B0B9 influenza a
26	34	81.0	238	2	Q7PM86	Q7PM86 anopheles g
27	34	81.0	238	2	Q9WHX7	Q9WHX7 influenza a
28	34	81.0	239	2	Q9WHX9	Q9WHX9 influenza a
29	34	81.0	239	2	Q9WHY1	Q9WHY1 influenza a
30	34	81.0	239	2	Q9WHY3	Q9WHY3 influenza a
31	34	81.0	239	2	Q9WHY4	Q9WHY4 influenza a

32	34	81.0	239	2	Q9WHY5	Q9WHY5 influenza a
33	34	81.0	239	2	Q9WHY6	Q9WHY6 influenza a
34	34	81.0	239	2	Q9WHY7	Q9WHY7 influenza a
35	34	81.0	239	2	Q9WHY8	Q9WHY8 influenza a
36	34	81.0	239	2	Q9WHY9	Q9WHY9 influenza a
37	34	81.0	249	2	Q67EA6	Q67EA6 influenza a
38	34	81.0	249	2	Q67EA8	Q67EA8 influenza a
39	34	81.0	249	2	Q67EB0	Q67EB0 influenza a
40	34	81.0	249	2	Q67EB2	Q67EB2 influenza a
41	34	81.0	249	2	Q67EB4	Q67EB4 influenza a
42	34	81.0	249	2	Q67EB6	Q67EB6 influenza a
43	34	81.0	249	2	Q67EB8	Q67EB8 influenza a
44	34	81.0	249	2	Q67EC0	Q67EC0 influenza a
45	34	81.0	249	2	Q67EC2	Q67EC2 influenza a

ALIGNMENTS

RESULT 1
Q8IYN3 PRELIMINARY; PRT; 1131 AA.
AC Q8IYN3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to PTPJL-associated RhoGAP 1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035564; AAK35564.1; -.
DR HSSP; Q98935; 1F7C.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001060; Cdc15 Fes CIP4.
DR InterPro; IPR000345; CytC Heme BS.
DR InterPro; IPR002219; DAG_Fg_bind.
DR InterPro; IPR000198; RhoGAP.
DR InterPro; IPR008936; RhoGAP.
DR Pfam; PF00130; Cl_1; 1.
DR Pfam; PF00620; RhoGAP; 1.
DR SMART; SM00109; Cl; 1.
DR SMART; SM00055; FCH; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN 1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50238; RHO_GAP; 1.
FT NON_TER 1
SQ SEQUENCE 1131 AA; 124081 MW; A606BIAA52014925 CRC64;

Query Match 100.0%; Score 42; DB 2; Length 1131;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRDDLEA 9
| | | | | | | | | |
Db 132 VLRDDLEA 140

RESULT 2
Q8HX84 PRELIMINARY; PRT; 1136 AA.
AC Q8HX84;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Minor histocompatibility antigen HA-1.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20549028; PubMed=11095984; DOI=10.1006/bbrc.2000.3880;
RA Kaminski W.E., Piehler A., Schmitz G.;
RT "Genomic organization of the human cholesterol-responsive ABC transporter ABCA7: tandem linkage with the minor histocompatibility antigen HA-1 gene.";
RT Biochem. Biophys. Res. Commun. 278:782-789 (2000).
RL EMBL; AF308045; AAN04658.1; JOINED.
DR EMBL; AF308046; AAN04658.1; JOINED.
DR EMBL; AF308047; AAN04658.1; JOINED.
DR EMBL; AF308048; AAN04658.1; JOINED.
DR EMBL; AF308049; AAN04658.1; JOINED.
DR EMBL; AF308050; AAN04658.1; JOINED.
DR EMBL; AF308051; AAN04658.1; JOINED.
DR EMBL; AF308052; AAN04658.1; JOINED.
DR EMBL; AF308053; AAN04658.1; JOINED.
DR EMBL; AF308054; AAN04658.1; JOINED.
DR EMBL; AF308055; AAN04658.1; JOINED.
DR EMBL; AF308056; AAN04658.1; JOINED.
DR EMBL; AF308057; AAN04658.1; JOINED.
DR EMBL; AF308058; AAN04658.1; JOINED.
DR EMBL; AF308059; AAN04658.1; JOINED.
DR EMBL; AF308060; AAN04658.1; JOINED.
DR EMBL; AF308061; AAN04658.1; JOINED.
DR EMBL; AF308062; AAN04658.1; JOINED.
DR EMBL; AF308063; AAN04658.1; JOINED.
DR EMBL; AF308064; AAN04658.1; JOINED.
DR EMBL; AF308065; AAN04658.1; JOINED.
DR EMBL; AF311102; AAN04658.1; JOINED.
DR HSP; Q98935; IFC.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001060; Cdc15 Fes CIP4.
DR InterPro; IPR000345; CytC heme BS.
DR InterPro; IPR002219; DAG PE-bind.
DR InterPro; IPR000198; RhoGAP.
DR Pfam; PF00130; Cl_1; 1.
DR SMART; SM00109; Cl; 1.
DR SMART; SM0055; FCH; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS00479; DAG PE BIND DOM 1; UNKNOWN 1.
DR PROSITE; PS0081; DAG PE BIND DOM 2; 1.
DR PROSITE; PS0238; RHO GAP; 1.
SQ SEQUENCE 1136 AA; 124613 MW; 9536787B3B1E16D CRC64;

Query Match 100.0%; Score 42; DB 2; Length 1136;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRDLDLEA 9
Db 137 VLRDLDLEA 145

RESULT 3
Q92619 ID Q92619 PRELIMINARY; PRT; 1165 AA.
AC Q92619;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DR KIAA0223 protein (Fragment).
GN Name=KIAA0223;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K., Ohira M., Kawarabayashi Y., Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI. The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by analysis of cDNA clones from cell line KG-1 and brain.";
RT DNA Res. 3:321-329 (1996).
RL EMBL; D86976; BAA13212.1; -.
DR PIR; D59433; D59433.
DR HSP; Q98935; IFC.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR Pfam; PF00130; Cl_1; 1.
DR SMART; SM00109; Cl; 1.
DR SMART; SM0055; FCH; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS00479; DAG PE BIND DOM 1; UNKNOWN 1.
DR PROSITE; PS0081; DAG PE BIND DOM 2; 1.
DR PROSITE; PS0238; RHO GAP; 1.
FT NON TER 1
SQ SEQUENCE 1165 AA; 127344 MW; 92EF768CAFD458C9 CRC64;

Query Match 100.0%; Score 42; DB 2; Length 1165;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRDLDLEA 9
Db 166 VLRDLDLEA 174

RESULT 4
Q8MVB1 ID Q8MVB1 PRELIMINARY; PRT; 376 AA.
AC Q8MVB1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative serine protease with signal anchor.
OS Ixodes scapularis (Black-legged tick) (Deer tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.
OX NCBI_TaxID=6945;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rhode Island; TISSUE=Salivary gland;
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M., Mather T.N., Ribeiro J.M.C.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AF483729; AAM93651.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001545; Gly hormoneB.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00261; GLYC HORMONE BETA_1; UNKNOWN 1.
DR PROSITE; PS00240; TRYPSIN DOM_1;
DR PROSITE; PS00134; TRYPSIN HIS; UNKNOWN 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 376 AA; 41434 MW; 64F403CAD505CD33 CRC64;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 06:44:47 ; Search time 90.5 Seconds
(without alignments)
38.462 Million cell updates/sec

Title: US-10-623-176-10

Perfect score: 42

Sequence: 1 VLKDDLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	2 AAW99197	AAW99197 Minor his
2	42	100.0	9	2 AAW97374	AAW97374 HA-1 R-al
3	42	100.0	9	8 ADH40334	Adh40334 Human min
4	42	100.0	13	2 AAW99198	AAW99198 DH cell K
5	42	100.0	13	2 AAW97414	AAW97414 KIAA0223
6	42	100.0	1136	5 ADI11793	Adi11793 Human NOV
7	42	100.0	1136	5 ADI16744	Adi16744 Human NOV
8	42	100.0	1136	8 ADH61299	Adh61299 INTSIG pr
9	42	100.0	1136	8 ADN42398	Adn42398 Human nov
10	42	100.0	1136	8 ADS10490	Adsl0490 Human the
11	42	100.0	1165	5 ADI17194	Adi17194 Human NOV
12	42	100.0	1165	8 ADP55035	Adp55035 Human PRO
13	42	100.0	1165	8 ADP24600	Adp24600 PRO polyp
14	38	90.5	4150	3 AAY92707	Aay92707 S. antibi
15	37	88.1	9	2 AAW99196	AAW99196 Minor his
16	37	88.1	9	2 AAW97375	AAW97375 HA-1 H-al
17	37	88.1	9	8 ADH40333	Adh40333 Human min
18	37	88.1	13	2 AAW99199	AAW99199 vR cell K
19	37	88.1	13	2 AAW97415	AAW97415 KIAA0223
20	37	88.1	1165	8 ADS11768	Adsl1768 Human the
21	37	88.1	1177	6 ABP71254	Abp71254 Chimeric
22	37	88.1	1186	6 ABP71253	Abp71253 Zebrafish
23	36	85.7	9	2 AAW99195	AAW99195 Minor his
24	36	85.7	9	2 AAW97572	AAW97572 T-cell ep
25	34	81.0	282	5 ABG93191	ABg93191 S. cerevi

26	34	81.0	324	4	AAU51225	AAU51225 Propionib
27	34	81.0	324	6	ABM47744	ABm47744 Propionib
28	34	81.0	343	8	ADN18393	Adn18393 Bacterial
29	34	81.0	486	6	ABU48977	ABu48977 Protein e
30	34	81.0	3519	3	AAU92709	AAy92709 S. antibi
31	34	81.0	3816	3	AAU92708	AAy92708 S. antibi
32	33	78.6	293	4	AAU35037	AAu35037 Enterococ
33	33	78.6	293	6	ABU29365	ABu29365 Protein e
34	33	78.6	328	4	AAG73861	AGg73861 Human col
35	33	78.6	472	7	ADC07956	Adc07956 Rice prot
36	33	78.6	486	7	ADC08309	Adc08309 Rice prot
37	33	78.6	491	7	ADC08289	Adc08289 Rice prot
38	33	78.6	738	8	ADS29025	AdS29025 Bacterial
39	33	78.6	822	6	ABP58031	ABp58031 Human EGF
40	33	78.6	822	7	ADN95677	ADn95677 Human EGF
41	33	78.6	925	4	AAE06696	AAe06696 Arabidops
42	33	78.6	1041	6	ABU34693	ABu34693 Protein e
43	33	78.6	1041	6	ABU36647	ABu36647 Protein e
44	33	78.6	1045	2	AAW57448	AAw57448 M. tuberc
45	33	78.6	1059	6	ABU35902	ABu35902 Protein e

ALIGNMENTS

RESULT 1

AAW99197

ID AAW99197 standard; peptide; 9 AA.

AC AAW99197;

DT 20-MAY-1999 (first entry)

XX Minor histocompatibility antigen HA-1 T-cell epitope #3.

XX Minor histocompatibility antigen; HA-1: T-cell epitope; immunological;

XX graft versus host disease; bone marrow transplant; leukaemia; vaccine;

XX diagnosis; aplastic anaemia; immune deficiency disease.

OS Homo sapiens.

PN WO9905174-A1.

PD 04-FEB-1999.

PF 23-JUL-1998; 98WO-NL000425.

PR 23-JUL-1997; 97EP-00202303.

XX (UYLE-) RIJKSUNIV LEIDEN.

XX Goulmy EAJM, Hunt DF, Engelhard VH;

XX WPI; 1999-153312/13.

XX A new minor histocompatibility antigen, HA-1 - useful to treat immune

XX diseases and prevent rejection and host versus graft disease in bone

XX marrow and organ transplantation.

XX Disclosure; Page 15; 47pp; English.

XX The present sequence represents a new peptide (PI) constituting a T-cell
XX epitope obtainable from the minor histocompatibility antigen HA-1. The
XX peptide is immunogenic and can be used as part of a vaccine. PI is used
XX as a medicine, to induce tolerance for transplants, prevent rejection
XX and/or graft versus host disease, or to treat (auto) immune diseases. In
XX particular it can be used with bone marrow transplantation, in the
XX treatment of severe aplastic anaemia, leukaemia, and immune deficiency
XX diseases

XX Sequence 9 AA;

Query Match

100.0%; Score 42; DB 2; Length 9;

```

Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRDALLEA 9
Db 1 VLRDALLEA 9

RESULT 2
AAW97374
ID AAW97374 standard; protein; 9 AA.
AC AAW97374;
XX
XX
XX 13-MAY-1999 (first entry)
XX
XX HA-1 R-allele sequence.
XX
XX Intron; minor histocompatibility antigen HA-1; typing allele; H allele;
XX R allele; polymorphic nucleotide; HA-1 typing; bone marrow transplant;
XX severe aplastic anaemia; leukaemia; immune deficiency disease; ss.
XX
XX Homo sapiens.
XX
XX WO9905313-A2.
XX
XX 04-FEB-1999.
XX
XX 23-JUL-1998; 98WO-EP004928.
XX
XX 23-JUL-1997; 97EP-00202303.
XX 02-JUN-1998; 98EP-00870125.
XX
XX (UYLE-) RIJKSUNIV LEIDEN.
XX
XX Goulmy E;
XX
XX WPI; 1999-142960/12.
XX
XX Typing minor histocompatibility antigen HA-1 - by amplifying and
XX identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g. detection
XX of genetic aberrances.
XX
XX Claim 13; Fig 5; 59pp; English.
XX
XX The present sequence represents part of the minor histocompatibility
XX antigen HA-1 R-allele. The specification describes methods for typing
XX alleles (preferably the H and R alleles) of the minor histocompatibility
XX antigen HA-1 in a sample, which comprise detecting polymorphic
XX nucleotides in the cDNA or genomic nucleic acids of the alleles. The
XX methods can be used for HA-1 typing for bone marrow transplants, severe
XX aplastic anaemia, leukaemia and immune deficiency diseases, as well as
XX detection of genetic aberrances. The probes and primers of the invention
XX can be used to screen for the HA-1 alleles. The HA-1 peptides can be used
XX anti-idiotypic B cells and/or T cells and antibodies
XX
XX Sequence 9 AA;

Query Match 100.0%; Score 42; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRDALLEA 9
Db 1 VLRDALLEA 9

RESULT 3
ADH40334
ID ADH40334 standard; peptide; 9 AA.
XX
XX ADH40334;
XX

```

```

DT 11-MAR-2004 (first entry)
XX
XX Human minor histocompatibility antigen HA-1 T cell epitope.
XX
XX human; cytostatic; vaccine; SNP profile; cancer; leukaemia;
XX minor histocompatibility antigen; mHag; T cell epitope.
XX
XX Homo sapiens.
XX
XX WO2003106692-A2.
XX
XX 24-DEC-2003.
XX
XX 13-JUN-2003; 2003WO-EP006251.
XX
XX 13-JUN-2002; 2002EP-00013423.
XX
XX (MERE ) MERCK PATENT GMBH.
XX
XX Strittmatter W, Moll H;
XX
XX WPI; 2004-082200/08.
XX
XX Providing allelic variant epitope of protein based on single nucleotide
XX polymorphism by defining target protein, screening database of protein,
XX identifying, selecting allelic variant protein, creating variant
XX epitopes.
XX
XX Disclosure; Page 82; 119pp; English.
XX
XX The invention relates to a novel method for providing epitopes of allelic
XX variants of antigenic proteins from specific species based on single
XX nucleotide polymorphism (SNP), by defining target protein/peptide or its
XX subset, screening database of DNA encoding target protein, identifying,
XX selecting allelic peptide/protein variants, expression product or its
XX fragment encoded by DNA sequence having SNP, creating variant epitopes,
XX selecting epitopes binding to MHC protein. A protein of the invention has
XX cytostatic activity, and may have a use in a vaccine. The method is
XX useful for generating a SNP profile of one or more individuals from a
XX given species by applying the method for several protein from the
XX individuals, where the SNP profile was related to disease, preferably
XX generating the SNP-related polymorphic profile. A method of the invention
XX is useful for transplanting haematopoietic stem cells from a donor to a
XX recipient and treating cancer, preferably leukaemia, and for determining
XX the progression, regression or onset of a treated disease. The present
XX sequence is used in the exemplification of the invention.
XX
XX Sequence 9 AA;

Query Match 100.0%; Score 42; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRDALLEA 9
Db 1 VLRDALLEA 9

RESULT 4
AAW99198
ID AAW99198 standard; peptide; 13 AA.
XX
XX AAW99198;
XX
XX 20-MAY-1999 (first entry)
XX
XX DH cell KIAA0223 protein sequence.
XX
XX Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
XX graft versus host disease; bone marrow transplant; leukaemia; vaccine;
XX diagnosis; aplastic anaemia; immune deficiency disease.
XX
XX

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:09:04 ; Search time 70 Seconds
(without alignments)
44.445 Million cell updates/sec

Title: US-10-623-176-10

Perfect score: 42

Sequence: 1 VLKDDLEA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 1465611

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	42	100.0	9	15	US-10-623-176-10
2	42	100.0	9	16	US-10-791-217-5
3	42	100.0	10	15	US-10-623-176-44
4	42	100.0	12	15	US-10-623-176-65
5	42	100.0	13	15	US-10-623-176-74
6	42	100.0	13	15	US-10-791-217-14
7	42	100.0	17	15	US-10-623-176-69
8	42	100.0	19	15	US-10-623-176-70
9	42	100.0	21	15	US-10-623-176-71
10	42	100.0	22	15	US-10-623-176-63
11	42	100.0	23	15	US-10-623-176-72
12	42	100.0	25	15	US-10-623-176-67
13	42	100.0	26	15	US-10-623-176-66
					Sequence 10, Appl
					Sequence 5, Appl
					Sequence 44, Appl
					Sequence 65, Appl
					Sequence 74, Appl
					Sequence 14, Appl
					Sequence 68, Appl
					Sequence 70, Appl
					Sequence 71, Appl
					Sequence 63, Appl
					Sequence 72, Appl
					Sequence 67, Appl

14	42	100.0	27	15	US-10-623-176-60	Sequence 60, Appl
15	42	100.0	28	15	US-10-623-176-61	Sequence 61, Appl
16	42	100.0	29	15	US-10-623-176-59	Sequence 59, Appl
17	42	100.0	1136	15	US-10-072-012-280	Sequence 280, App
18	42	100.0	1136	15	US-10-072-012-729	Sequence 729, App
19	42	100.0	1165	15	US-10-072-012-730	Sequence 730, App
20	38	90.5	9	15	US-10-623-176-42	Sequence 42, Appl
21	38	90.5	9	15	US-10-623-176-46	Sequence 46, Appl
22	38	90.5	4150	10	US-09-808-880-2	Sequence 2, Appl
23	37	88.1	9	15	US-10-623-176-2	Sequence 2, Appl
24	37	88.1	9	16	US-10-791-217-2	Sequence 1, Appl
25	37	88.1	9	17	US-10-861-335-1	Sequence 1, Appl
26	37	88.1	10	15	US-10-623-176-43	Sequence 43, Appl
27	37	88.1	13	15	US-10-623-176-76	Sequence 76, Appl
28	37	88.1	13	16	US-10-791-217-16	Sequence 16, Appl
29	37	88.1	17	15	US-10-623-176-57	Sequence 57, Appl
30	37	88.1	22	15	US-10-623-176-53	Sequence 53, Appl
31	37	88.1	25	15	US-10-623-176-55	Sequence 55, Appl
32	37	88.1	27	15	US-10-623-176-51	Sequence 51, Appl
33	37	88.1	29	15	US-10-623-176-49	Sequence 49, Appl
34	37	88.1	1177	14	US-10-193-692-4	Sequence 4, Appl
35	37	88.1	1186	14	US-10-193-692-2	Sequence 2, Appl
36	36	85.7	9	15	US-10-623-176-1	Sequence 1, Appl
37	36	85.7	9	16	US-10-791-217-1	Sequence 1, Appl
38	36	85.7	10	15	US-10-623-176-24	Sequence 24, Appl
39	36	85.7	17	15	US-10-623-176-33	Sequence 33, Appl
40	34	81.0	9	15	US-10-623-176-48	Sequence 48, Appl
41	34	81.0	282	16	US-10-451-467A-340	Sequence 340, App
42	34	81.0	343	15	US-10-369-493-1046	Sequence 1046, Ap
43	34	81.0	407	16	US-10-437-963-177253	Sequence 177253,
44	34	81.0	408	16	US-10-437-963-114620	Sequence 114620,
45	34	81.0	486	15	US-10-282-122A-76901	Sequence 76901, A

ALIGNMENTS

RESULT 1
US-10-623-176-10
; Sequence 10, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Els A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(9)
US-10-623-176-10

Query Match 100.0%; Score 42; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRRDLLEA 9
| | | | |
Db 1 VLRRDLLEA 9

RESULT 2

US-10-791-217-5
; Sequence 5, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; APPLICANT: Hunt, Donald F.
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-4285US
; CURRENT APPLICATION NUMBER: US/10/791,217
; CURRENT FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: obtained from KIAA0223 partial complementary DNA
US-10-791-217-5

Query Match 100.0%; Score 42; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRRDLLEA 9
| | | | |
Db 1 VLRRDLLEA 9

RESULT 3

US-10-623-176-44
; Sequence 44, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; APPLICANT: Hunt, Donald F.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1) .. (10)
US-10-623-176-44

Query Match 100.0%; Score 42; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRRDLLEA 9
| | | | |
Db 1 VLRRDLLEA 9

RESULT 4

US-10-623-176-65
; Sequence 65, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; APPLICANT: Hunt, Donald F.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; NAME/KEY: SITE
; LOCATION: (1) .. (12)
US-10-623-176-65

Query Match 100.0%; Score 42; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRRDLLEA 9
| | | | |
Db 2 VLRRDLLEA 10

RESULT 5

US-10-623-176-74
; Sequence 74, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; APPLICANT: Hunt, Donald F.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
US-10-623-176-74

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 06:54:48 ; Search time 24 Seconds
(without alignments)
27.993 Million cell updates/sec

Title: US-10-623-176-10

Perfect score: 42

Sequence: 1 VLRDDLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	4	US-09-269-250E-18
2	42	100.0	13	4	US-09-269-250E-26
3	42	100.0	26	4	US-09-269-250E-36
4	38	90.5	4150	3	US-09-428-517-2
5	37	88.1	9	4	US-09-269-250E-20
6	37	88.1	13	4	US-09-269-250E-28
7	36	85.7	9	4	US-09-269-250E-29
8	34	81.0	3519	3	US-09-428-517-4
9	34	81.0	3816	3	US-09-428-517-3
10	33	78.6	324	4	US-09-154-750A-89
11	33	78.6	576	4	US-09-949-016-8057
12	33	78.6	822	4	US-09-886-319A-34
13	33	78.6	1045	1	US-08-452-083-2
14	32	76.2	126	4	US-09-134-000C-6541
15	32	76.2	192	4	US-08-198-452A-404
16	32	76.2	234	4	US-09-270-767-33603
17	32	76.2	373	4	US-09-328-352-7186
18	32	76.2	427	4	US-09-252-991A-30099
19	32	76.2	507	4	US-09-438-185A-387
20	32	76.2	613	4	US-09-489-039A-10059
21	32	76.2	1196	4	US-09-252-991A-18261
22	31	73.8	251	5	PCT-US95-12357A-1
23	31	73.8	252	4	US-09-506-286B-2
24	31	73.8	252	4	US-09-506-286B-5
25	31	73.8	252	4	US-09-762-861B-2
26	31	73.8	252	4	US-09-762-861B-5
27	31	73.8	252	4	US-10-065-133A-2

28	31	73.8	252	4	US-10-065-133A-5	Sequence 5, Appli
29	31	73.8	252	4	US-10-434-811A-2	Sequence 2, Appli
30	31	73.8	252	4	US-10-434-811A-5	Sequence 5, Appli
31	31	73.8	379	4	US-05-540-236-3485	Sequence 3485, Ap
32	31	73.8	383	4	US-09-248-796A-14833	Sequence 14833, A
33	31	73.8	433	4	US-09-902-540-12839	Sequence 12839, A
34	31	73.8	508	2	US-08-861-464-10	Sequence 10, Appl
35	31	73.8	508	2	US-08-396-001-10	Sequence 10, Appl
36	31	73.8	508	3	US-09-323-433A-10	Sequence 10, Appl
37	31	73.8	508	4	US-09-826-752-10	Sequence 10, Appl
38	31	73.8	591	4	US-09-134-000C-6729	Sequence 6729, Ap
39	31	73.8	599	3	US-08-556-419-22	Sequence 22, Appl
40	31	73.8	599	4	US-09-252-991A-20462	Sequence 20462, A
41	31	73.8	629	3	US-08-556-419-23	Sequence 23, Appl
42	31	73.8	635	4	US-09-252-991A-31013	Sequence 31013, A
43	31	73.8	648	4	US-09-919-039-42	Sequence 42, Appl
44	31	73.8	692	4	US-09-248-796A-18612	Sequence 18612, A
45	31	73.8	820	4	US-09-328-352-5766	Sequence 5766, Ap

ALIGNMENTS

RESULT 1

US-09-269-250E-18
; Sequence 18, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Exon Fragments
US-09-269-250E-18

Query Match 100.0%; Score 42; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VLRDDLEA	9
Db	1	VLRDDLEA	9

RESULT 2

US-09-269-250E-26
; Sequence 26, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PCR Product
US-09-269-250E-26

Query Match 100.0%; Score 42; DB 4; Length 13;

```
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRRDLEA 9
Db 3 VLRRDLEA 11

RESULT 3
US-09-269-250E-36
; Sequence 36, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PCR Product
US-09-269-250E-36

Query Match 100.0%; Score 42; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRRDLEA 9
Db 3 VLRRDLEA 11

RESULT 4
US-09-428-517-2
; Sequence 2, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
; OTHER INFORMATION: Oleandolide PKS
US-09-428-517-2

Query Match 90.5%; Score 38; DB 3; Length 4150;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LRDDLEA 9
Db 1781 LRDDLEA 1788
```

```
RESULT 5
US-09-269-250E-20
; Sequence 20, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Exon Fragments
US-09-269-250E-20

Query Match 88.1%; Score 37; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLRRDLEA 9
Db 1 VLRRDLEA 9

RESULT 6
US-09-269-250E-28
; Sequence 28, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PCR Product
US-09-269-250E-28

Query Match 88.1%; Score 37; DB 4; Length 13;
Best Local Similarity 88.9%; Pred. No. 0.26;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLRRDLEA 9
Db 3 VLRRDLEA 11

RESULT 7
US-09-269-250E-29
; Sequence 29, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 9
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:30:19 ; Search time 16.25 Seconds
(without alignments)
53.289 Million cell updates/sec

Title: US-10-623-176-2

Perfect score: 45

Sequence: 1 VLHDDLLEA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 791

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	40.0	7	2	146868
2	17	37.8	9	2	S55696
3	16	35.6	7	2	P00663
4	16	35.6	7	2	S68004
5	16	35.6	8	2	PC4131
6	15	33.3	6	2	S78764
7	15	33.3	7	2	A59489
8	14	31.1	7	2	S20446
9	14	31.1	8	2	XGHUEU
10	14	31.1	9	2	A12872
11	14	31.1	9	2	S10920
12	14	31.1	9	2	A60427
13	13	28.9	9	2	A61328
14	13	28.9	9	2	PH0942
15	12	26.7	4	2	I40697
16	12	26.7	5	2	PT0679
17	12	26.7	5	2	PT0601
18	12	26.7	6	2	B35640
19	12	26.7	6	2	PT0533
20	12	26.7	7	2	A34026
21	12	26.7	7	2	B39040
22	12	26.7	7	2	PT0628
23	12	26.7	7	2	PT0722
24	12	26.7	7	2	PT0576
25	12	26.7	8	2	PT0368
26	12	26.7	8	2	PN0043
27	12	26.7	8	2	PT0557
28	12	26.7	9	2	PM0002
29	12	26.7	9	2	S65913

RESULT 1
146868
alpha-myosin heavy chain - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 09-Jul-2004
C;Accession: I46868
R;Friedman, D.J.; Umeda, P.K.; Sinha, A.M.; Hsu, H.
Proc. Natl. Acad. Sci. U.S.A. 81, 3044-3048, 1984
A;Title: Characterization of genomic clones specifying rabbit alpha- and beta-ventricular
A;Reference number: I46868; MUID:84221901; PMID:6328491
A;Accession: I46868
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-7 <FRI>
A;Cross-references: UNIPROT:Q28742; GB:K01698; NID:g165538; PIDN:AAA31415.1; PID:g165539

Query Match 40.0%; Score 18; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LHDD 5
Db 3 MHDE 6

RESULT 2
S55696
phosphoenolpyruvate carboxykinase - Trypanosoma brucei
C;Species: Trypanosoma brucei
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55696
R;Hunt, M.; Koehler, P.
Biochim. Biophys. Acta 1249, 15-22, 1995
A;Title: Purification and characterization of phosphoenolpyruvate carboxykinase from Try
A;Reference number: S55696; MUID:95284106; PMID:7766679
A;Accession: S55696
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <HUN>
A;Cross-references: UNIPROT:Q7M3S5

Query Match 37.8%; Score 17; DB 2; Length 9;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLHDDL 6
Db 3 ITHKNL 8

RESULT 3
PQ0663

ALIGNMENTS

```

membrane protein - porcine epidemic diarrhea virus (isolate Belgian CV777) (fragment)
C;Species: porcine epidemic diarrhea virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 08-Oct-1999
C;Accession: PQ0663
R;Bridgen, A.; Duarte, M.; Tobler, K.; Laude, H.; Ackermann, M.
J. Gen. Virol. 74, 1795-1804, 1993
A;Title: Sequence determination of the nucleocapsid protein gene of the porcine epidemic
issile gastroenteritis virus.
A;Reference number: JQ2191; MUID:93389433; PMID:8397280
A;Accession: PQ0663
A;Molecule type: mRNA
A;Residues: 1-7 <BRI>
A;Cross-references: GS:Z14976; NID:G311650; PIDN:CAA78699.1; PID:G584083
C;Comment: This virus is coronavirus related to human coronavirus 229E.
C;Keywords: membrane protein

Query Match      35.6%; Score 16; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLH 3
   |||
Db 3 VLH 5

RESULT 4
S68004
hucolin, 75K chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S68004
R;Edgar, P.F.
FEBS Lett. 375, 159-161, 1995
A;Title: Hucolin, a new corticosteroid-binding protein from human plasma with structural
A;Reference number: S68004; MUID:96087107; PMID:7498469
A;Accession: S68004
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <EDG>

Query Match      35.6%; Score 16; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6
   |||
Db 4 DDL 6

RESULT 5
PC4131
hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)
C;Species: Pseudomonas aeruginosa
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: PC4131
R;Kawasaki, S.; Arai, H.; Igarashi, Y.; Kodama, T.
Gene 167, 87-91, 1995
A;Title: Sequencing and characterization of the downstream region of the genes encoding
y for biosynthesis of heme dl.
A;Reference number: JC4552; MUID:96144254; PMID:8566817
A;Accession: PC4131
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-8 <XAW>
A;Cross-references: UNIPROT:P95412; DBJ:D50473; NID:gl217594
A;Note: this ORF is not annotated in GenBank entry PSENIIRC, release 113.0

Query Match      35.6%; Score 16; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6
   |||

```

```

Db 2 DDL 4

RESULT 6
S78764
ribosomal protein MRP-S23, mitochondrial - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: S78764
R;Graack, H.R.
submitted to the Protein Sequence Database, July 1999
A;Reference number: S78760
A;Accession: S78764
A;Molecule type: protein
A;Residues: 1-6 <GRA>
C;Keywords: mitochondrial
F;1-6/Product: ribosomal protein MRP-S23 (fragment) #status experimental <MAT>

Query Match      33.3%; Score 15; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LHDD 5
   |||
Db 1 LHVD 4

RESULT 7
A59489
protein kinase C inhibitor - rat (fragment)
C;Species: Rattus norvegicus
C;Date: 25-Aug-2003 #sequence_revision 25-Aug-2003 #text_change 25-Aug-2003
C;Accession: A59489
R;Negoro, M.
submitted to the Protein Sequence Database, June 2003
A;Description: Purification of PKCI from rat liver.
A;Reference number: A59489
A;Accession: A59489
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <NEG>
A;Experimental source: strain Wistar, liver
A;Note: p-Hydroxyacetophenone-Sepharose binding protein

Query Match      33.3%; Score 15; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLHDD 5
   : |||
Db 2 IFEDD 6

RESULT 8
S20446
elastase - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 24-Jul-1997
C;Accession: S20446
R;Kessler, E.; Saffrin, M.; Peretz, M.; Burstein, Y.
FEBS Lett. 299, 291-293, 1992
A;Title: Identification of cleavage sites involved in proteolytic processing of Pseudomor
A;Reference number: S20446; MUID:92183956; PMID:1544509
A;Accession: S20446
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <KES>

Query Match      31.1%; Score 14; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLE 8

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:11:54 ; Search time 77.5 Seconds
(without alignments)
59.467 Million cell updates/sec

Title: US-10-623-176-2

Perfect score: 45

Sequence: 1 VLHDDLLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1455

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	19	42.2	9	2	Q7R8X5	Q7R8X5 plasmodium
2	19	42.2	9	2	Q47556	Q47556 escherichia
3	18	40.0	7	2	Q28742	Q28742 tryptolagus
4	17	37.8	9	2	Q7M3S5	Q7M3S5 trypanosoma
5	16	35.6	7	2	Q99182	Q99182 gnatholebia
6	16	35.6	8	2	Q726G0	Q726G0 homo sapien
7	16	35.6	9	1	FAR8 MACRS	P83281 macrobrachi
8	16	35.6	9	2	Q70SM2	Q70SM2 homo sapien
9	16	35.6	9	2	Q8LPT5	Q8LPT5 zea mays (m
10	16	35.6	9	2	P82568	P82568 streptococc
11	15	33.3	8	2	Q15900	Q15900 homo sapien
12	15	33.3	9	2	Q9XJN0	Q9XJN0 bacterioph
13	14	31.1	6	1	TRPI_PSEPU	P36414 pseudomonas
14	14	31.1	8	1	GLUR_HUMAN	P02729 homo sapien
15	14	31.1	8	2	Q9BFA7	Q9BFA7 macroscelid
16	14	31.1	8	2	P72279	P72279 rhodococcus
17	14	31.1	8	2	Q93SR0	Q93SR0 staphylococ
18	14	31.1	9	1	TALI_PICJA	P17440 pichia jadi
19	14	31.1	9	2	Q7M4R5	Q7M4R5 homo sapien
20	14	31.1	9	2	Q7M471	Q7M471 vespa orien
21	14	31.1	9	2	Q6Q7G0	Q6Q7G0 rana ridibu
22	14	31.1	9	2	Q9FSZ2	Q9FSZ2 cicier ariet
23	14	31.1	9	2	Q8CG39	Q8CG39 rattus norv
24	14	31.1	9	2	Q9QZA8	Q9QZA8 mus musculu
25	14	31.1	9	2	Q8UTD7	Q8UTD7 human immun
26	14	31.1	9	2	Q85710	Q85710 rous sarcom
27	13	28.9	7	2	Q66205	Q66205 transmissib
28	13	28.9	8	2	Q9HDS4	Q9HDS4 aspergillus
29	13	28.9	8	2	Q15889	Q15889 homo sapien
30	13	28.9	8	2	Q7M390	Q7M390 balaenopter
31	13	28.9	8	2	Q37854	Q37854 bacterioph

32	13	28.9	8	2	Q6JC68	Q6JC68 glycine max
33	13	28.9	8	2	Q7XB03	Q7XB03 zea mays (m
34	13	28.9	8	2	Q51594	Q51594 plaemid col
35	13	28.9	9	2	Q6UVK2	Q6UVK2 malus domes
36	13	28.9	9	2	Q7X8P7	Q7X8P7 zea mays (m
37	13	28.9	9	2	Q712A6	Q712A6 sinorthizobi
38	13	28.9	9	2	Q8G1Z6	Q8G1Z6 lactobacill
39	12	26.7	5	1	BIOA_CITFR	P13071 citrobacter
40	12	26.7	6	1	TMOF_SARBU	P41495 sarcophaga
41	12	26.7	7	1	UF03_MOUSE	P38641 mus musculu
42	12	26.7	7	2	O49223	O49223 glycine max
43	12	26.7	8	1	VTIN_ALOVR	P83233 aloe vera (
44	12	26.7	8	2	Q15894	Q15894 homo sapien
45	12	26.7	8	2	Q16468	Q16468 homo sapien

ALIGNMENTS

RESULT 1

Q7R8X5 PRELIMINARY; PRT; 9 AA.

AC Q7R8X5; 26, Created)

DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)

DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)

DE Hypothetical protein.

GN Name=PY07095;

OS Plasmodium yoelii yoelii.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=73239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=17XNL;

RX PubMed=12368865; DOI=10.1038/nature01099;

RA Carlton J.M., Angiolini S.V., Suh B.B., Kooij T.W., Pertea M., Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L., Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L., Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V., Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M., Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A., Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B., van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R., Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J., Carucci D.J.;

RA "Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii.";

RT Nature 419:512-519(2002)

RL

CC -!- CAUTION: The sequence shown here is derived from an preliminary data.

CC EMBL/GenBank/DDJB whole genome shotgun (WGS) entry which is preliminary data.

DR EMBL; AABL01002528; EAA19452.1; -.

KW Hypothetical protein.

SQ SEQUENCE 9 AA; 1013 MW; 4684D447244441E7 CRC64;

Query Match 42.2%; Score 19; DB 2; Length 9;

Best Local Similarity 50.0%; Pred. No. 1.6e+06;

Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDDL 6

Db 1 MLHNNL 6

RESULT 2

Q47556 PRELIMINARY; PRT; 9 AA.

ID Q47556;

AC Q47556;

DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)

DE Aspartate transcarbamoylase regulatory chain (Fragment).

GN Name=pyri;

```
OS Escherichia coli K12.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83333;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82275057; PubMed=7051000;
RA Pauza C.D., Karels M.J., Navre M., Schachman H.K.;
RT "Genes encoding Escherichia coli aspartate transcarbamoylase: the
RT pyrB-pyri operon.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:4020-4024 (1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83195078; PubMed=6302686;
RA Hoover T.A., Roof W.D., Foltermann K.F., O'Donovan G.A., Bencini D.A.,
RA Wild J.R.;
RT "Nucleotide sequence of the structural gene (pyrB) that encodes the
RT catalytic polypeptide of aspartate transcarbamoylase of Escherichia
RT coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:2462-2466 (1983).
DR EMBL; J01670; AAA24475.1; -.
FT NON TER 9
FT SEQUENCE 9 AA; 1085 MW; 99EFD723344AA1F1 CRC64;

Query Match 42.2%; Score 19; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.6e+06;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 HDDLLE 8
Db 3 HDNKLQ 8

RESULT 3
Q28742 ID Q28742 PRELIMINARY; PRT; 7 AA.
AC Q28742;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alpha-myosin heavy chain (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84221901; PubMed=6328491;
RA Friedman D.J., Umeda P.K., Sinha A.M., Hsu H.-J., Jokovic S.,
RA Rabinowitz M.;
RT "Characterization of genomic clones specifying rabbit alpha- and beta-
RT ventricular myosin heavy chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:3044-3048 (1984).
DR EMBL; K01698; AAA31415.1; -.
DR PIR; I46868; I46868.
FT NON TER 1
FT SEQUENCE 7 AA; 916 MW; 6B1B1AA1E69326B0 CRC64;

Query Match 40.0%; Score 18; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.6e+06;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LHDD 5
Db 3 MHDE 6

RESULT 4
Q7M3S5 ID Q7M3S5 PRELIMINARY; PRT; 9 AA.
AC Q7M3S5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphoenolpyruvate carboxykinase.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE.
RX MEDLINE=95284106; PubMed=7766679; DOI=10.1016/0167-4838(95)00061-X;
RA Hunt M., Koehler P.;
RT "Purification and characterization of phosphoenolpyruvate
RT carboxykinase from Trypanosoma brucei.";
RL Biochim. Biophys. Acta 1249:15-22 (1995).
DR PIR; S55696; S55696.
FT NON TER 9
FT SEQUENCE 9 AA; 1063 MW; 35F2244331805047 CRC64;

Query Match 37.8%; Score 17; DB 2; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.6e+06;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLHDDL 6
Db 3 IIHKNL 8

RESULT 5
Q09182 ID Q09182 PRELIMINARY; PRT; 7 AA.
AC Q09182;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome oxidase I (Fragment).
GN Name=COI;
OS Gnatholebias zonatus.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Aplocheilidae; Rivulinae; Gnatholebias.
OX NCBI_TaxID=135316;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20072928; PubMed=10603257; DOI=10.1006/mpev.1999.0656;
RA Murphy W.J., Thomson J.E., Collier G.E.;
RT "Phylogeny of the Neotropical killifish family Rivulidae
RT (Cyprinodontiformes, Aplocheiloidei) inferred from mitochondrial DNA
RT sequences.";
RL Mol. Phylogenet. Evol. 13:289-301 (1999).
DR EMBL; AF002591; AAD01074.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 1
FT SEQUENCE 7 AA; 899 MW; 672721F6CB572030 CRC64;

Query Match 35.6%; Score 16; DB 2; Length 7;
Best Local Similarity 42.9%; Pred. No. 1.6e+06;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VLHDDL 7
Db 1 ILYOHLL 7

RESULT 6
Q7Z6G0 ID Q7Z6G0 PRELIMINARY; PRT; 8 AA.
AC Q7Z6G0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fumarate hydratase (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:10:14 ; Search time 85.5 Seconds
(without alignments)
40.712 Million cell updates/sec

Title: US-10-623-176-2

Perfect score: 45

Sequence: 1 VLHDDLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 330156

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq l6Dec04.*

1: geneseqp1980s.*

2: geneseqp1990s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	9	2 AAW99196	AAW99196 Minor his
2	45	100.0	9	2 AAW97375	AAW97375 HA-1 H-al
3	45	100.0	9	8 ADH40333	ADH40333 Human min
4	37	82.2	9	2 AAW99197	AAW99197 Minor his
5	37	82.2	9	2 AAW97374	AAW97374 HA-1 R-al
6	37	82.2	9	8 ADH40334	ADH40334 Human min
7	36	80.0	9	2 AAW99195	AAW99195 Minor his
8	36	80.0	9	2 AAW97572	AAW97572 T-cell ep
9	30	66.7	9	2 AAW97373	AAW97373 Peptide e
10	28	62.2	9	2 AAY10122	AAY10122 T cell ep
11	28	62.2	9	5 ABG79805	ABG79805 MHC class
12	28	62.2	9	8 ADK68732	ADK68732 Epitope 1
13	28	62.2	9	8 ADK05291	ADK05291 Hepatitis
14	28	62.2	9	8 ADK05293	ADK05293 Hepatitis
15	28	62.2	9	8 ADQ10530	ADQ10530 Cercopith
16	27	60.0	9	6 ABV20115	ABV20115 MHC bindi
17	26	57.8	6	6 ABU97233	ABU97233 Enzyme pe
18	26	57.8	6	6 ABU97232	ABU97232 Enzyme pe
19	26	57.8	6	6 ABU97234	ABU97234 Enzyme pe
20	26	57.8	6	6 ABU97231	ABU97231 Enzyme pe
21	26	57.8	7	8 ADM96262	ADM96262 Human ser
22	26	57.8	9	5 AAE31275	AAE31275 Human mag
23	26	57.8	9	8 ADM96261	ADM96261 Human ser
24	26	57.8	9	8 ADM96258	ADM96258 Human ser
25	25	55.6	6	6 ABU97229	ABU97229 Enzyme pe

26	25	55.6	6	6 ABU97230	ABU97230 Enzyme pe
27	25	55.6	9	4 AAG79144	AAG79144 Amino ter
28	25	55.6	9	4 AAB84492	AAB84492 Peptide f
29	25	55.6	9	5 AAE26631	AAE26631 Yeast GPA
30	25	55.6	9	7 ADC25889	ADC25889 Yeast GPA
31	24	53.3	6	2 AAR29358	AAR29358 Endotheli
32	24	53.3	6	2 AAR69140	AAR69140 Endotheli
33	24	53.3	7	2 AAW33386	AAW33386 Altered I
34	24	53.3	7	4 AAB70589	AAB70589 Human imm
35	24	53.3	7	5 AAO21066	AAO21066 Isomerise
36	24	53.3	7	5 AAE28095	AAE28095 Human imm
37	24	53.3	9	2 AAR43735	AAR43735 MHC Class
38	24	53.3	9	2 AAR43734	AAR43734 MHC Class
39	24	53.3	9	2 AAW39656	AAW39656 HPV16/18
40	24	53.3	9	2 AAW39657	AAW39657 HPV16/18
41	24	53.3	9	2 AAW54759	AAW54759 Peptide f
42	24	53.3	9	2 AAY10561	AAY10561 HLA Class
43	24	53.3	9	5 ABG80244	ABG80244 MHC class
44	24	53.3	9	8 ADK69184	ADK69184 Epitope 1
45	24	53.3	9	8 ADK05290	ADK05290 Hepatitis

ALIGNMENTS

RESULT 1

AAW99196

ID AAW99196 standard; peptide; 9 AA.

XX AC AAW99196;

XX DT 20-MAY-1999 (first entry)

XX DE Minor histocompatibility antigen HA-1 T-cell epitope #2.

XX KW Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;

XX KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;

XX KW diagnosis; aplastic anaemia; immune deficiency disease.

XX OS Homo sapiens.

XX PN WO9905174-A1.

XX PD 04-FEB-1999.

XX PF 23-JUL-1998; 98WO-NL000425.

XX PR 23-JUL-1997; 97EP-00202303.

XX FA (UYLE-) RIJKSUNIV LEIDEN.

XX PI Goulmy EAJM, Hunt DF, Engelhard VH;

XX DR WPI; 1999-153312/13.

XX PT A new minor histocompatibility antigen, HA-1 - useful to treat immune

XX PT diseases and prevent rejection and host versus graft disease in bone

XX XX marrow and organ transplantation.

XX PS Claim 3; Page 32; 47pp; English.

XX CC The present sequence represents a new peptide (PI) constituting a T-cell

XX CC epitope obtainable from the minor histocompatibility antigen HA-1. The

XX CC peptide is immunogenic and can be used as part of a vaccine. PI is used

XX CC as a medicine, to induce tolerance for transplants, prevent rejection

XX CC and/or graft versus host disease, or to treat (auto) immune diseases. In

XX CC particular it can be used with bone marrow transplantation, in the

XX CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency

XX CC diseases

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 2; Length 9;

```

Best Local Similarity 100.0%; Pred. No. 1.8e+06; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

Qy 1 VLHDDLLEA 9
Db 1 VLHDDLLEA 9

RESULT 2
AAW97375
ID AAW97375 standard; protein; 9 AA.
XX
AC AAW97375;
XX
DT 13-MAY-1999 (first entry)
XX
DE HA-1 H-allele sequence.
XX
KW Intron; minor histocompatibility antigen HA-1; typing allele; H allele;
KW R allele; polymorphic nucleotide; HA-1 typing; bone marrow transplant;
KW severe aplastic anaemia; leukaemia; immune deficiency disease; ss.
XX
OS Homo sapiens.
XX
PN WO9905313-A2.
XX
PD 04-FEB-1999.
XX
PF 23-JUL-1998; 98WO-EP004928.
XX
PR 23-JUL-1997; 97EP-00202303.
XX
PR 02-JUN-1998; 98EP-00870125.
XX
PA (UYLE-) RIJKSUNIV LEIDEN.
XX
PI Goulmy E;
XX
DR WPI; 1999-142960/12.
XX
XX Typing minor histocompatibility antigen HA-1 - by amplifying and
PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g. detection
PT of genetic aberrances.
XX
PS Claim 18; Fig 5; 59pp; English.
XX
CC The present sequence represents part of the minor histocompatibility
CC antigen HA-1 H-allele. The specification describes methods for typing
CC alleles (preferably the H and R alleles) of the minor histocompatibility
CC antigen HA-1 in a sample, which comprise detecting polymorphic
CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The
CC methods can be used for HA-1 typing for bone marrow transplants, severe
CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as
CC detection of genetic aberrances. The probes and primers of the invention
CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be used
CC anti-idiotypic B cells and/or T cells and antibodies
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
Db 1 VLHDDLLEA 9

RESULT 3
ADH40333
ID ADH40333 standard; peptide; 9 AA.
XX
AC ADH40333;
XX
XX

```

```

DT 11-MAR-2004 (first entry)
XX
DE Human minor histocompatibility antigen HA-1 T cell epitope.
XX
KW human; cytostatic; vaccine; SNP profile; cancer; leukaemia;
KW minor histocompatibility antigen; mRag; T cell epitope.
XX
OS Homo sapiens.
XX
PN WO2003106692-A2.
XX
PD 24-DEC-2003.
XX
PF 13-JUN-2003; 2003WO-EP006251.
XX
PR 13-JUN-2002; 2002EP-00013423.
XX
PA (MERE ) MERCK PATENT GMBH.
XX
PI Strittmatter W, Moll H;
XX
DR WPI; 2004-082200/08.
XX
XX Providing allelic variant epitope of protein based on single nucleotide
PT polymorphism by defining target protein, screening database of protein,
PT identifying, selecting allelic variant protein, creating variant
PT epitopes.
XX
PS Disclosure; Page 82; 119pp; English.
XX
CC The invention relates to a novel method for providing epitopes of allelic
CC variants of antigenic proteins from specific species based on single
CC nucleotide polymorphism (SNP), by defining target protein/peptide or its
CC subset, screening database of DNA encoding target protein, identifying,
CC selecting allelic peptide/protein variants, expression product or its
CC fragment encoded by DNA sequence having SNP, creating variant epitopes,
CC selecting epitopes binding to MHC protein. A protein of the invention has
CC cytostatic activity, and may have a use in a vaccine. The method is
CC useful for generating a SNP profile of one or more individuals from a
CC given species by applying the method for several protein from the
CC individuals, where the SNP profile was related to disease, preferably
CC generating the SNP-related polymorphic profile. A method of the invention
CC is useful for transplanting haematopoietic stem cells from a donor to a
CC recipient and treating cancer, preferably leukaemia, and for determining
CC the progression, regression or onset of a treated disease. The present
CC sequence is used in the exemplification of the invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 45; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDDLLEA 9
Db 1 VLHDDLLEA 9

RESULT 4
AAW99197
ID AAW99197 standard; peptide; 9 AA.
XX
AC AAW99197;
XX
DT 20-MAY-1999 (first entry)
XX
DE Minor histocompatibility antigen HA-1 T-cell epitope #3.
XX
KW Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;
KW diagnosis; aplastic anaemia; immune deficiency disease.
XX
XX

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:48:11 ; Search time 67 Seconds
(without alignments)
46.435 Million cell updates/sec

Title: US-10-623-176-2

Perfect score: 45

Sequence: 1 VLHDDLLLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 142046

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*

10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*

11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*

12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*

13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*

14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*

15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*

16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*

17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*

18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*

19: /cgn2_6/ptodata/1/pubpaa/US10G_PUBCOMB.pep.*

20: /cgn2_6/ptodata/1/pubpaa/US10H_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	9	15	US-10-623-176-2
2	45	100.0	9	16	US-10-791-217-2
3	45	100.0	9	17	US-10-861-335-1
4	41	91.1	9	15	US-10-623-176-41
5	41	91.1	9	15	US-10-623-176-45
6	37	82.2	9	15	US-10-623-176-10
7	37	82.2	9	15	US-10-623-176-47
8	37	82.2	9	16	US-10-791-217-5
9	36	80.0	9	15	US-10-623-176-1
10	36	80.0	9	15	US-10-623-176-40
11	36	80.0	9	16	US-10-791-217-1
12	33	73.3	9	15	US-10-623-176-42
13	33	73.3	9	15	US-10-623-176-46

14	32	71.1	9	15	US-10-623-176-4	Sequence 4, Appli
15	30	66.7	9	15	US-10-623-176-14	Sequence 14, Appli
16	30	66.7	9	16	US-10-791-217-4	Sequence 4, Appli
17	29	64.4	9	15	US-10-623-176-48	Sequence 48, Appli
18	28.5	63.3	8	15	US-10-623-176-15	Sequence 15, Appli
19	28	62.2	9	15	US-10-623-176-23	Sequence 23, Appli
20	28	62.2	9	15	US-10-623-176-38	Sequence 38, Appli
21	28	62.2	9	16	US-10-777-053-95	Sequence 95, Appli
22	27	60.0	9	17	US-10-705-459-280	Sequence 280, App
23	26	57.8	6	14	US-10-166-225A-163	Sequence 163, App
24	26	57.8	6	14	US-10-166-225A-164	Sequence 164, App
25	26	57.8	6	14	US-10-166-225A-165	Sequence 165, App
26	26	57.8	6	14	US-10-166-225A-166	Sequence 166, App
27	26	57.8	6	17	US-10-825-026-48	Sequence 48, Appli
28	25	55.6	6	14	US-10-166-225A-161	Sequence 161, App
29	25	55.6	6	14	US-10-166-225A-162	Sequence 162, App
30	25	55.6	7	9	US-09-803-951-5	Sequence 5, Appli
31	25	55.6	9	9	US-09-747-774A-3	Sequence 3, Appli
32	25	55.6	9	10	US-09-953-354-3	Sequence 3, Appli
33	25	55.6	9	14	US-10-267-074-14	Sequence 14, Appli
34	25	55.6	9	14	US-10-267-074-16	Sequence 16, Appli
35	25	55.6	9	17	US-10-752-478-3	Sequence 3, Appli
36	24	53.3	7	14	US-10-020-354-86	Sequence 86, Appli
37	24	53.3	8	17	US-10-808-187-462	Sequence 462, Appli
38	24	53.3	9	15	US-10-623-176-5	Sequence 5, Appli
39	24	53.3	9	16	US-10-777-053-547	Sequence 547, App
40	23	51.1	8	9	US-09-982-172-74	Sequence 74, Appli
41	23	51.1	9	9	US-09-834-765-132	Sequence 132, App
42	23	51.1	9	9	US-09-834-765-349	Sequence 349, App
43	23	51.1	9	15	US-10-245-871-785	Sequence 785, App
44	23	51.1	9	15	US-10-245-871-809	Sequence 809, App
45	23	51.1	9	15	US-10-253-286-785	Sequence 785, App

ALIGNMENTS

RESULT 1

US-10-623-176-2

; Sequence 2, Application US/10623176

; Publication No. US20040092446A1

; GENERAL INFORMATION:

; APPLICANT: Goulmy, Els A.J.M.

; APPLICANT: Hunt, Donald F.

; TITLE OF INVENTION: HA-1 epitopes and uses thereof

; FILE REFERENCE: 2183-6047US

; CURRENT APPLICATION NUMBER: US/10/623.176

; CURRENT FILING DATE: 2003-07-18

; PRIOR APPLICATION NUMBER: 09/489,760

; PRIOR FILING DATE: 2000-01-21

; PRIOR APPLICATION NUMBER: EP 97202303.0

; PRIOR FILING DATE: 1997-07-23

; PRIOR APPLICATION NUMBER: PCT/NL98/00424

; PRIOR FILING DATE: 1998-07-23

; PRIOR APPLICATION NUMBER: JP 2000-504165

; PRIOR FILING DATE: 2000-01-24

; NUMBER OF SEQ ID NOS: 77

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide

; FEATURE:

; NAME/KEY: SITE -

; LOCATION: (1)..(9)

US-10-623-176-2

Query Match 100.0%; Score 45; DB 15; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.3e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDDLLEA 9
| | | | |
Db 1 VLHDDLLEA 9

RESULT 2
US-10-791-217-2
; Sequence 2, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-4285US
; CURRENT APPLICATION NUMBER: US/10/791,217
; PRIOR FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: obtained from histocompatibility antigen
US-10-791-217-2

Query Match 100.0%; Score 45; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDDLLEA 9
| | | | |
Db 1 VLHDDLLEA 9

RESULT 3
US-10-861-335-1
; Sequence 1, Application US/10861335
; Publication No. US20050031612A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; TITLE OF INVENTION: Minor histocompatibility antigen HA-1: target antigen for immunot
; FILE REFERENCE: 2183-6479US
; CURRENT APPLICATION NUMBER: US/10/861,335
; CURRENT FILING DATE: 2004-06-04
; PRIOR APPLICATION NUMBER: PCT/NL02/00791
; PRIOR FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: EP 01204704.9
; PRIOR FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HA-1 peptide
US-10-861-335-1

Query Match 100.0%; Score 45; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDDLLEA 9
| | | | |
Db 1 VLHDDLLEA 9

RESULT 4
US-10-623-176-41
; Sequence 41, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Els A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; NAME/KEY: SITE
; LOCATION: (1)..(9)
US-10-623-176-41

Query Match 91.1%; Score 41; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDDLLE 8
| | | | |
Db 2 VLHDDLLE 9

RESULT 5
US-10-623-176-45
; Sequence 45, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Els A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; FEATURE:

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:32:15 ; Search time 23 Seconds
(without alignments)
29.211 Million cell updates/sec

Title: US-10-623-176-2

Perfect score: 45

Sequence: 1 VLHDDLLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 99282

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	9	4	US-09-269-250E-20
2	37	82.2	9	4	US-09-269-250E-18
3	36	80.0	9	4	US-09-269-250E-29
4	28.5	63.3	8	4	US-09-269-250E-38
5	25	55.6	7	3	US-09-025-819-5
6	25	55.6	7	4	US-09-808-126-5
7	25	55.6	7	4	US-09-803-951-5
8	25	55.6	9	3	US-08-582-333A-3
9	25	55.6	9	3	US-09-305-923A-7
10	23	51.1	6	3	US-09-217-609A-8
11	23	51.1	6	3	US-08-873-235B-8
12	23	51.1	9	3	US-08-159-339A-370
13	22	48.9	5	3	US-08-811-463-30
14	22	48.9	5	4	US-09-933-497B-30
15	22	48.9	7	1	US-08-208-036-7
16	22	48.9	7	1	US-08-208-036-9
17	22	48.9	7	1	US-08-428-823-7
18	22	48.9	7	1	US-08-428-823-9
19	22	48.9	7	3	US-08-556-419-14
20	22	48.9	7	3	US-09-173-941-82
21	22	48.9	7	4	US-09-494-190-82
22	22	48.9	8	1	US-08-403-378B-13
23	22	48.9	9	3	US-09-217-609A-11
24	22	48.9	9	3	US-08-873-235B-11
25	21	46.7	6	2	US-08-459-568-20
26	21	46.7	6	2	US-08-399-411-20
27	21	46.7	6	3	US-08-516-859A-20

```

28 21 46.7 6 3 US-09-586-472-20 Sequence 20, Appl
29 21 46.7 6 4 US-09-528-706-20 Sequence 20, Appl
30 21 46.7 8 2 US-08-459-568-75 Sequence 75, Appl
31 21 46.7 8 2 US-08-399-411-75 Sequence 75, Appl
32 21 46.7 8 3 US-08-516-859A-75 Sequence 75, Appl
33 21 46.7 8 3 US-09-419-826-1 Sequence 1, Appl
34 21 46.7 8 3 US-09-419-826-37 Sequence 37, Appl
35 21 46.7 8 3 US-09-586-472-75 Sequence 75, Appl
36 21 46.7 8 4 US-09-528-706-75 Sequence 75, Appl
37 21 46.7 9 1 US-07-671-757-37 Sequence 30, Appl
38 21 46.7 9 1 US-08-343-602-2 Sequence 2, Appl
39 21 46.7 9 2 US-08-459-568-72 Sequence 72, Appl
40 21 46.7 9 2 US-08-459-568-78 Sequence 78, Appl
41 21 46.7 9 2 US-08-399-411-72 Sequence 72, Appl
42 21 46.7 9 2 US-08-399-411-78 Sequence 78, Appl
43 21 46.7 9 3 US-08-516-859A-72 Sequence 72, Appl
44 21 46.7 9 3 US-08-516-859A-78 Sequence 78, Appl
45 21 46.7 9 3 US-09-502-600-76 Sequence 76, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-269-250E-20
; Sequence 20, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Exon Fragments
US-09-269-250E-20

```

```

Query Match 100.0%; Score 45; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 VLHDDLLEA 9
Db 1 VLHDDLLEA 9

```

```

RESULT 2
US-09-269-250E-18
; Sequence 18, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Exon Fragments
US-09-269-250E-18

```

```

Query Match 82.2%; Score 37; DB 4; Length 9;

```

```
Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLHDDLLLEA 9
   || || || || || || || ||
Db 1 VLKDDLLLEA 9

RESULT 3
US-09-269-250E-29
; Sequence 29, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3)-(3)
; OTHER INFORMATION: Xaa represents a histidine (H) or an arginine (R) residue
US-09-269-250E-29

Query Match 80.0%; Score 36; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLHDDLLLEA 9
   || || || || || || || ||
Db 1 VLKDDLLLEA 9

RESULT 4
US-09-269-250E-38
; Sequence 38, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 8
; TYPE: PRT
; ORGANISM: HUMAN
US-09-269-250E-38

Query Match 63.3%; Score 28.5; DB 4; Length 8;
Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 VLHDDLLLEA 9
   || || || || || || || ||
Db 1 VLKDDLLLEA 8

RESULT 5
US-09-025-819-5
; Sequence 5, Application US/09025819
; Patent No. 623097
; GENERAL INFORMATION:
; APPLICANT: Obata, Shusei
; APPLICANT: Nishino, Tokuzo
```

```
; APPLICANT: Koyama, Tanetoshi
; APPLICANT: Sato, Yoshihiro
; TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KENYON & KENYON
; STREET: 1500 K Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,819
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 251675
; FILING DATE: 17-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Khalilian, Hourii
; REGISTRATION NUMBER: 39,546
; REFERENCE/DOCKET NUMBER: 10235/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-220-4200
; TELEFAX: 202-220-4201
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-025-819-5

Query Match 55.6%; Score 25; DB 3; Length 7;
Best Local Similarity 80.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHDD 5
   : || || ||
Db 3 LLHDD 7

RESULT 6
US-09-808-126-5
; Sequence 5, Application US/09808126
; Patent No. 6410280
; GENERAL INFORMATION:
; APPLICANT: Obata, Shusei
; APPLICANT: Nishino, Tokuzo
; APPLICANT: Koyama, Tanetoshi
; APPLICANT: Sato, Yoshihiro
; TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KENYON & KENYON
; STREET: 1500 K Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/808,126
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:30:19 ; Search time 16.25 Seconds
(without alignments)
53.289 Million cell updates/sec

Title: US-10-623-176-10

Perfect score: 42

Sequence: 1 VLRDILLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 791

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	38.1	7	2 S68004	hucolin, 75K chain
2	16	38.1	8	2 PC4131	hypothetical prote
3	15	35.7	7	2 A59489	protein kinase C i
4	14	33.3	7	2 S20446	elastase - Pseudom
5	14	33.3	9	2 A60427	macrophage cytotox
6	13	31.0	9	2 PH0942	T-cell receptor be
7	12	28.6	4	2 I40697	biotin A - Citroba
8	12	28.6	5	2 PT0679	T-cell receptor be
9	12	28.6	5	2 PT0601	T-cell receptor be
10	12	28.6	6	2 B35640	cerebellar degener
11	12	28.6	6	2 PT0533	T-cell receptor be
12	12	28.6	7	2 A34026	acetylcholinestera
13	12	28.6	7	2 B39040	calsequestrin, fas
14	12	28.6	7	2 PT0628	T-cell receptor be
15	12	28.6	7	2 PT0722	T-cell receptor be
16	12	28.6	7	2 PT0576	T-cell receptor be
17	12	28.6	8	2 PT0557	T-cell receptor be
18	12	28.6	8	2 A61328	trypsin (EC 3.4.21
19	12	28.6	9	2 PH0108	late G1-69 protein
20	12	28.6	9	2 PT0562	T-cell receptor be
21	12	28.6	9	2 B30572	T-cell receptor be
22	12	28.6	9	2 A39841	sucrose 3-glucosyl
23	12	28.6	9	2 A37027	macrophage chemota
24	11	26.2	5	2 D60274	major protein anti
25	11	26.2	6	2 T11779	phosphoglycerate t
26	11	26.2	7	2 S25266	p11E protein - Esc
27	11	26.2	7	2 PT0602	T-cell receptor be
28	11	26.2	8	2 S22428	chitin-binding pro
29	11	26.2	8	2 B33099	158K exoantigen -

30	11	26.2	8	2 S69165	ferredoxin a2 - Ja
31	11	26.2	9	2 B28455	conopressin S - co
32	11	26.2	9	2 PW0002	chlorophyll a/b-bi
33	11	26.2	9	2 B39841	dextranucrase (EC
34	11	26.2	9	2 S66419	tetrameric protein
35	11	26.2	9	2 PS0253	glycine cleavage s
36	11	26.2	9	2 A12872	transaldolase (EC
37	11	26.2	9	2 B57444	neuropeptide Grb-A
38	11	26.2	9	2 C57444	neuropeptide Grb-A
39	11	26.2	9	2 A44787	calliFMRFamide 10
40	10	23.8	4	2 A48360	gamma subunit of p
41	10	23.8	4	2 A26209	protein-glutamine
42	10	23.8	4	2 A35779	neuropeptide Antho
43	10	23.8	6	2 B44510	hypothetical prote
44	10	23.8	6	2 H48394	glycoprotein compo
45	10	23.8	7	2 S78024	ribosomal protein

ALIGNMENTS

RESULT 1

S68004
hucolin, 75K chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S68004
R;Edgar, P.F.
FEBS Lett. 375, 159-161, 1995
A;Title: Hucolin, a new corticosteroid-binding protein from human plasma with structural
A;Reference number: S68004; MUID:96087107; PMID:7498469
A;Accession: S68004
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <EDG>

Query Match 38.1%; Score 16; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	4	DDL 6
Db	4	DDL 6

RESULT 2

PC4131
hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)
C;Species: Pseudomonas aeruginosa
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: PC4131
R;Kawasaki, S.; Arai, H.; Igarashi, Y.; Kodana, T.
Gene 167, 87-91, 1995
A;Title: Sequencing and characterization of the downstream region of the genes encoding r
y for biosynthesis of heme d1.
A;Reference number: JC4552; MUID:96144254; PMID:8566817
A;Accession: PC4131
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-8 <KAW>
A;Cross-references: UNIPROT:P95412; DDBJ:D50473; NID:gl217594
A;Note: this ORF is not annotated in Genbank entry PSENIIRC, release 113.0

Query Match 38.1%; Score 16; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	4	DDL 6
Db	2	DDL 4

RESULT 3

A59489
protein kinase C inhibitor - rat (fragment)
C;Species: Rattus norvegicus
C;Date: 25-Aug-2003 #sequence_revision 25-Aug-2003 #text_change 25-Aug-2003
C;Accession: A59489
R;Negoro, M.
submitted to the Protein Sequence Database, June 2003
A;Description: Purification of PKCI from rat liver.
A;Reference number: A59489
A;Accession: A59489
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <NEG>
A;Experimental source: strain Wistar, liver
A;Note: p-Hydroxyacetophenone-Sepharose binding protein

Query Match 35.7%; Score 15; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VLRDD 5
: ||
Db 2 IFEDD 6

RESULT 4
S20446
elastase - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 24-Jul-1997
C;Accession: S20446
R;Kessler, E.; Safran, M.; Peretz, M.; Burstein, Y.
FEBS Lett. 299, 291-293, 1992
A;Title: Identification of cleavage sites involved in proteolytic processing of Pseudom
A;Reference number: S20446; MUID:92183956; PMID:1544509
A;Accession: S20446
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <KES>

Query Match 33.3%; Score 14; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 DLLE 8
: ||:
Db 3 DLID 6

RESULT 5
A60427
macrophage cytotoxicity-inducing factor, 29K - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C;Accession: A60427
R;Jones, C.M.; Prince, C.A.; Williams, J.S.
Exp. Hematol. 19, 704-709, 1991
A;Title: Purification and amino acid analysis of a human macrophage cytotoxicity-inducin
A;Reference number: A60427; MUID:91372335; PMID:1909970
A;Accession: A60427
A;Molecule type: protein
A;Residues: 1-9 <JON>
A;Cross-references: UNIPROT:Q7M4R5
A;Note: the sequence from the text on page 706 is inconsistent with that from page 708 i
C;Keywords: cytokine

Query Match 33.3%; Score 14; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLRD 4
: ||
Db 4 VLED 7

RESULT 6

PH0942

T-cell receptor beta chain V-D-J region (clone 13) - rat (fragment)

C;Species: Rattus norvegicus (Norway rat)

C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997

C;Accession: PH0942

R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.

J. Exp. Med. 174, 1467-1476, 1991

A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic

A;Reference number: PH0891; MUID:92078857; PMID:1836012

A;Accession: PH0942

A;Molecule type: mRNA

A;Residues: 1-9 <GOL>

A;Experimental source: complete Freund's adjuvant-immunized lymph node

A;Note: the authors translated the codon TGC for residue 2 as Ala

C;Keywords: T-cell receptor

Query Match 31.0%; Score 13; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LLE 8

: |||

Db 5 LLE 7

RESULT 7

I40697

biotin A - Citrobacter freundii (fragment)

C;Species: Citrobacter freundii

C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004

C;Accession: I40697

R;Shiuan, D.; Campbell, A.

Gene 67, 203-211, 1988

A;Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citrobacter;

A;Reference number: I40697; MUID:89006280; PMID:2971595

A;Accession: I40697

A;Status: preliminary; translated from GB/EMBL/DBD

A;Molecule type: DNA

A;Residues: 1-4 <RES>

A;Cross-references: UNIPROT:P13071; GB:M21922; NID:gl44434

Query Match 28.6%; Score 12; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DD 5

: ||

Db 3 DD 4

RESULT 8

PT0679

T-cell receptor beta chain V-D-J region - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C;Accession: PT0679; PT0708

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0679

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-5 <PEE>

A;Experimental source: day 18 fetal thymus, strain BALB/c, 154-2J

A;Accession: PT0708

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-5 <PE2>

A;Experimental source: newborn thymus, strain BALB/c, 161-2B

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:11:54 ; Search time 77.5 Seconds
(without alignments)
59.467 Million cell updates/sec

Title: US-10-623-176-10
Perfect score: 42
Sequence: 1 VLRDLEA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1455

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	42.9	8	2	Q7XB03	Q7xb03 zea mays (m
2	18	42.9	9	2	Q7X8P7	Q7x8p7 zea mays (m
3	16	38.1	9	2	Q70S42	Q70sm2 homo sapien
4	16	38.1	9	2	Q7KYP6	Q7kyp6 homo sapien
5	16	38.1	9	2	Q8LPT5	Q8lpt5 zea mays (m
6	16	38.1	9	2	P82568	P82568 streptococ
7	15	35.7	8	2	P72279	P72279 rhodococcus
8	15	35.7	8	2	Q80H91	Q80h91 newcastle d
9	15	35.7	9	2	Q9UE36	Q9ue26 homo sapien
10	15	35.7	9	2	Q9XJN0	Q9xjn0 bacterioph
11	15	35.7	9	2	P83539	P83539 lactobacill
12	14	33.3	8	2	Q8KPY4	Q8kpy4 microcystis
13	14	33.3	8	2	Q51594	Q51594 plasmid col
14	14	33.3	8	2	Q9ZIE9	Q9zie9 neisseria m
15	14	33.3	9	2	Q7M4R5	Q7m4r5 homo sapien
16	14	33.3	9	2	Q6Q7G0	Q6q7g0 rana ridibu
17	14	33.3	9	2	Q9FS22	Q9fsz2 cicier arlet
18	14	33.3	9	2	Q9QZ48	Q9qza8 mus musculu
19	14	33.3	9	2	Q8UTD7	Q8utd7 human immun
20	14	33.3	9	2	Q85710	Q85710 rous sarcom
21	13	31.0	7	2	Q15897	Q15897 homo sapien
22	13	31.0	8	2	Q9HDS4	Q9hds4 aspergillus
23	13	31.0	8	2	G6JC68	G6jc68 glycine max
24	13	31.0	8	2	Q93SR0	Q93sr0 staphylococ
25	13	31.0	8	2	Q9QVJ8	Q9qvj8 mus sp. mep
26	13	31.0	8	2	O89965	O89965 polyomaviru
27	13	31.0	8	2	Q6PUD5	Q6pud5 simian viru
28	13	31.0	8	2	Q6PUD7	Q6pud7 simian viru
29	13	31.0	8	2	O6PUD9	O6pud9 simian viru
30	13	31.0	8	2	Q6PUEL	Q6puel simian viru
31	13	31.0	8	2	Q6QXS9	Q6qxs9 simian viru

32 13 31.0 8 2 Q64IX4 Q64ix4 fundulus he
33 13 31.0 9 1 IFR_RHOVI IFR_rhoVI
34 13 31.0 9 2 Q7RY89 Q7ry89 neurospora
35 13 31.0 9 2 Q6LCV2 Q6lcv2 homo sapien
36 13 31.0 9 2 Q6UVK2 Q6uvk2 malus domes
37 13 31.0 9 2 Q9IBM8 Q9ibm8 simian viru
38 13 31.0 9 2 Q9PYK1 Q9pyk1 simian viru
39 13 31.0 9 2 Q64IX3 Q64ix3 fundulus he
40 12 28.6 5 1 BIOA_CITFR BIOa_citFR
41 12 28.6 7 2 O99182 O99182 gnatholebia
42 12 28.6 7 2 O49223 O49223 glycine max
43 12 28.6 8 1 VTIN_ALOVR P83233 aloe vera (
44 12 28.6 8 2 O6U7E2 O6u7r2 cryptococcu
45 12 28.6 8 2 Q7Z4N9 Q7z4n9 homo sapien

ALIGNMENTS

RESULT 1
Q7XB03 PRELIMINARY; PRT; 8 AA.
ID Q7XB03;
AC Q7XB03;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Phytoene synthase 2 (Fragment).
GN Name=psy2;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y-3;
RX MEDLINE=22779048; PubMed=12897253;
RA Palaisa K.A., Morgante M., Williams M., Rafalski A.;
RT "Contrasting effects of selection on sequence diversity and linkage
RT disequilibrium at two phytoene synthase loci";
RL Plant Cell 15:1795-1806(2003).
DR EMBL: AY300612; AAP55351.1; .
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 915 MW; 5D1862CAB4072054 CRC64;
Query Match 42.9%; Score 18; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.6e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VLRD 4
Db : : : :
2 ILRD 5

RESULT 2
Q7X8P7 PRELIMINARY; PRT; 9 AA.
ID Q7X8P7;
AC Q7X8P7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Phytoene synthase 2 (Fragment).
GN Name=psy2;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y-13, and W-16;
RX MEDLINE=22779048; PubMed=12897253;

RA Palaia K.A., Morgante M., Williams M., Rafalski A.;
 RT "Contrasting effects of selection on sequence diversity and linkage
 RL disequilibrium at two phytoene synthase loci";
 DR Plant Cell 15:1795-1806(2003).
 DR EMBL; AY300641; AAP55379.1; -;
 DR EMBL; AY300665; AAP55403.1; -;
 FT NON TER 1
 FT NON TER 9
 SQ SEQUENCE 9 AA; 1016 MW; 1D1C62CAB4072054 CRC64;

Query Match 42.9%; Score 18; DB 2; Length 9;
 Best Local Similarity 75.0%; Pred. No. 1.6e+06;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRD 4
 Db 3 ILRD 6

RESULT 3
 Q70SM2 PRELIMINARY; PRT; 9 AA.
 AC Q70SM2
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypermethylated in cancer 1 (Fragment).
 GN Name=HIC1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RA Pinte S., Guerdel C., Deltour S., Godwin A.K., Leprince D.;
 RT "Identification of a second G-C-rich promoter conserved in the human,
 RT murine and rat tumor suppressor genes HIC1";
 RL Oncogene 23:4023-4031(2004).
 DR EMBL; AJ550616; CAD79467.1; -;
 FT NON TER 9
 SQ SEQUENCE 9 AA; 964 MW; 5B5E6DDB1681AAA7 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;
 Best Local Similarity 60.0%; Pred. No. 1.6e+06;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DLLEA 9
 Db 3 DTMEA 7

RESULT 4
 Q7KYP6 PRELIMINARY; PRT; 9 AA.
 AC Q7KYP6
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Mi.V glycoprotein (exon P3B) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Miltnerberger class V;
 RX MEDLINE=90005483; PubMed=2792104;
 RA Vignal A., Rahuel C., El-Maliki B., London J., Le Vankim C.,
 RA Blanchard C.D.;
 RT "Molecular Analysis of glycophorin A and B gene structure and
 RT expression . . .";
 DR Eur. J. Biochem. 184:337-344(1989).

RA SEQUENCE FROM N.A.
 RC TISSUE=Miltnerberger class V;
 RA Vignal A., Rahuel C., El-Maliki B., London J., Le Vankim C.,
 RA Blanchard C.D.;
 RT "Molecular Analysis of glycophorin A and B gene structure and
 RT expression . . .";
 DR Eur. J. Biochem. 184:0-0(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Miltnerberger class V;
 RA Carttron J.P.;
 RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X15824; CAA33822.1; -;
 FT NON TER 9
 SQ SEQUENCE 9 AA; 1073 MW; 50D145B7244AB403 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;
 Best Local Similarity 75.0%; Pred. No. 1.6e+06;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RDDL 6
 Db 3 RDNL 6

RESULT 5
 Q8LPT5 PRELIMINARY; PRT; 9 AA.
 AC Q8LPT5
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Beta-expansin-like protein (Fragment).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 OC NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ching A.S., Caldwell K.S., Jung M., Dolan M., Smith O.S., Tingey S.,
 RA Morgante M., Rafalski J.A.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY094310; AM21836.1; -;
 FT NON TER 1
 SQ SEQUENCE 9 AA; 977 MW; 5C05B2D2CB1AAA3 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;
 Best Local Similarity 33.3%; Pred. No. 1.6e+06;
 Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DDLLEA 9
 Db 4 DEVVDA 9

RESULT 6
 P82568 PRELIMINARY; PRT; 9 AA.
 AC P82568
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Unknown protein from 2D-page (Fragment).
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC STRAIN=JRS4;
 RA Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,

GenCore version 5.1.6 -
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:10:14 ; Search time 85.5 Seconds
(without alignments)

40.712 Million cell updates/sec

Title: US-10-623-176-10

Perfect score: 42

Sequence: 1 VLRDLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 330156

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

1: geneseqp1980s.*

2: geneseqp1990s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	2 AAW99197	AAW99197 Minor his
2	42	100.0	9	2 AAW97374	AAW97374 HA-1 R-al
3	42	100.0	9	8 ADH40334	ADH40334 Human min
4	37	88.1	9	2 AAW99196	AAW99196 Minor his
5	37	88.1	9	2 AAW97375	AAW97375 HA-1 H-al
6	37	88.1	9	8 ADH40333	ADH40333 Human min
7	36	85.7	9	2 AAW99195	AAW99195 Minor his
8	36	85.7	9	2 AAW99572	AAW99572 T-cell ep
9	27	64.3	9	2 AAW47427	AAW47427 Prenyl di
10	25	59.5	9	5 AAW71428	AAW71428 Human MHC
11	24	57.1	7	4 AAB81550	AAB81550 Zif268 zi
12	24	57.1	9	7 ADM18256	ADM18256 C trachom
13	24	57.1	9	7 ADM18218	ADM18218 Chlamydia
14	23	54.8	9	5 AAU71211	AAU71211 Human MHC
15	23	54.8	9	5 ABR19981	ABR19981 Human can
16	23	54.8	9	6 ABR19426	ABR19426 Human can
17	23	54.8	9	6 ABR18988	ABR18988 Human can
18	23	54.8	9	6 ABR19578	ABR19578 Human can
19	23	54.8	9	6 ABR19174	ABR19174 Human can
20	23	54.8	9	6 ABR19780	ABR19780 Human can
21	22	52.4	9	2 AAW97373	AAW97373 Peptide e
22	22	52.4	9	8 ADF72091	ADF72091 Rabbit pe
23	21	50.0	6	3 AAB12035	AAB12035 Peptide #
24	21	50.0	6	5 AAU80810	AAU80810 Rat RD-in
25	21	50.0	6	7 ADE84629	ADE84629 Mammalia

26	21	50.0	7	5 ABP48411	ABP48411 Zinc fing
27	21	50.0	7	5 ABP48414	ABP48414 Zinc fing
28	21	50.0	7	5 ABP48419	ABP48419 Zinc fing
29	21	50.0	7	5 ABG79371	ABG79371 CXCR zinc
30	21	50.0	7	6 ABP96179	ABP96179 Zinc fing
31	21	50.0	7	7 ADA62240	ADA62240 Zinc fing
32	21	50.0	7	7 ADA62237	ADA62237 Zinc fing
33	21	50.0	7	7 ADA62245	ADA62245 Zinc fing
34	21	50.0	7	7 ADM20449	ADM20449 Synthetic
35	21	50.0	7	7 ADM20444	ADM20444 Synthetic
36	21	50.0	7	7 ADM20441	ADM20441 Synthetic
37	21	50.0	7	8 ADF94279	ADF94279 Human cel
38	21	50.0	8	2 AAW26557	AAW26557 Soluble p
39	21	50.0	8	3 AAB12089	AAB12089 Ad7 cel p
40	21	50.0	8	4 AAB46568	AAB46568 Integrin
41	21	50.0	8	4 AAB46535	AAB46535 Integrin
42	21	50.0	8	6 ABP99038	ABP99038 ErbB2 cel
43	21	50.0	8	7 ADE84662	ADE84662 Adenoviru
44	21	50.0	9	3 AAY80162	AAY80162 HLA-A2 re
45	21	50.0	9	3 AAB12092	AAB12092 Ad40 cel

ALIGNMENTS

RESULT 1

AAW99197

ID AAW99197 standard; peptide; 9 AA.

XX AAW99197;

AC AAW99197;

XX 20-MAY-1999 (first entry)

DT 20-MAY-1999 (first entry)

XX Minor histocompatibility antigen HA-1 T-cell epitope #3.

DE Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;

XX Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;

KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;

KW diagnosis; aplastic anaemia; immune deficiency disease.

XX Homo sapiens.

OS Homo sapiens.

XX WO9905174-A1.

PN WO9905174-A1.

XX 04-FEB-1999.

PD 04-FEB-1999.

XX 23-JUL-1998; 98WO-NL000425.

PF 23-JUL-1998; 98WO-NL000425.

XX 23-JUL-1997; 97EP-00202303.

PR 23-JUL-1997; 97EP-00202303.

XX (UYLE-) RIJKSUNIV LEIDEN.

PA (UYLE-) RIJKSUNIV LEIDEN.

XX Goulmy EAJM, Hunt DF, Engelhard VH;

PI Goulmy EAJM, Hunt DF, Engelhard VH;

XX WPI; 1999-153312/13.

DR WPI; 1999-153312/13.

XX A new minor histocompatibility antigen, HA-1 - useful to treat immune

PT diseases and prevent rejection and host versus graft disease in bone

PT marrow and organ transplantation.

XX Disclosure; Page 15; 47pp; English.

PS Disclosure; Page 15; 47pp; English.

XX The present sequence represents a new peptide (PI) constituting a T-cell

CC epitope obtainable from the minor histocompatibility antigen HA-1. The

CC peptide is immunogenic and can be used as part of a vaccine. PI is used

CC as a medicine, to induce tolerance for transplants, prevent rejection

CC and/or graft versus host disease, or to treat (auto) immune diseases. In

CC particular it can be used with bone marrow transplantation, in the

CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency

XX diseases

XX Sequence 9 AA;

SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRRDLLEA 9
Db 1 VLRRDLLEA 9

RESULT 2
AAW97374
ID AAW97374 standard; protein; 9 AA.
XX AC
XX AAW97374;
XX
DT 13-MAY-1999 (first entry)
XX
DE HA-1 R-allele sequence.
XX
KW Intron; minor histocompatibility antigen HA-1; typing allele; H allele;
KW R allele; polymorphic nucleotide; HA-1 typing; bone marrow transplant;
KW severe aplastic anaemia; leukaemia; immune deficiency disease; ss.
XX
OS Homo sapiens.
XX
XX WO9905313-A2.
XX
XX 04-FEB-1999.
XX
XX 23-JUL-1998; 98WO-EP004928.
XX
XX 23-JUL-1997; 97EP-00202303.
PR 02-JUN-1998; 98EP-00870125.
XX
XX (UYLE-) RIJXSUNIV LEIDEN.
PA
XX Goulmy E;
PI
XX WPI; 1999-142960/12.
XX
XX Typing minor histocompatibility antigen HA-1 - by amplifying and
PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g. detection
PT of genetic aberrances.
XX
XX Claim 13; Fig 5; 59pp; English.
XX
XX The present sequence represents part of the minor histocompatibility
CC antigen HA-1 R-allele. The specification describes methods for typing
CC alleles (preferably the H and R alleles) of the minor histocompatibility
CC antigen HA-1 in a sample, which comprise detecting polymorphic
CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The
CC methods can be used for HA-1 typing for bone marrow transplants, severe
CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as
CC detection of genetic aberrances. The probes and primers of the invention
CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be used
CC anti-idiotypic B cells and/or T cells and antibodies
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRRDLLEA 9
Db 1 VLRRDLLEA 9

RESULT 3
ADH40334
ID ADH40334 standard; peptide; 9 AA.
XX AC
XX ADH40334;
XX

DT 11-MAR-2004 (first entry)
XX
DE Human minor histocompatibility antigen HA-1 T cell epitope.
XX
KW human; cytostatic; vaccine; SNP profile; cancer; leukaemia;
KW minor histocompatibility antigen; mHAg; T cell epitope.
XX
OS Homo sapiens.
XX
XX WO2003106692-A2.
XX
XX 24-DEC-2003.
XX
XX 13-JUN-2003; 2003WO-EP006251.
XX
XX 13-JUN-2002; 2002EP-00013423.
PR
XX (MERE) MERCK PATENT GMBH.
PA
XX Strittmatter W, Moll H;
XX
XX WPI; 2004-082200/08.
XX
XX Providing allelic variant epitope of protein based on single nucleotide
PT polymorphism by defining target protein, screening database of protein,
PT identifying, selecting allelic variant protein, creating variant
PT epitopes.
XX
XX Disclosure; Page 82; 119pp; English.
XX
XX The invention relates to a novel method for providing epitopes of allelic
CC variants of antigenic proteins from specific species based on single
CC nucleotide polymorphism (SNP), by defining target protein/peptide or its
CC subset, screening database of DNA encoding target protein, identifying,
CC selecting allelic peptide/protein variants, expression product or its
CC segment encoded by DNA sequence having SNP, creating variant epitopes,
CC selecting epitopes binding to MHC protein. A protein of the invention has
CC cytostatic activity, and may have a use in a vaccine. The method is
CC useful for generating a SNP profile of one or more individuals from a
CC given species by applying the method for several protein from the
CC individuals, where the SNP profile was related to disease, preferably
CC cancer. This is useful for diagnosing a disease in an individual by
CC generating the SNP-related polymorphic profile. A method of the invention
CC is useful for transplanting haematopoietic stem cells from a donor to a
CC recipient and treating cancer, preferably leukaemia, and for determining
CC the progression, regression or onset of a treated disease. The present
CC sequence is used in the exemplification of the invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRRDLLEA 9
Db 1 VLRRDLLEA 9

RESULT 4
AAW99196
ID AAW99196 standard; peptide; 9 AA.
XX
XX AAW99196;
XX
XX 20-MAY-1999 (first entry)
DT
XX
DE Minor histocompatibility antigen HA-1 T-cell epitope #2.
XX
KW Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;
KW diagnosis; aplastic anaemia; immune deficiency disease.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:48:11 ; Search time 67 Seconds
(without alignments)
46.435 Million cell updates/sec

Title: US-10-623-176-10

Perfect score: 42

Sequence: 1 VLRDDLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 142046

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	9	15	US-10-623-176-10
2	42	100.0	9	16	US-10-791-217-5
3	38	90.5	9	15	US-10-623-176-42
4	38	90.5	9	15	US-10-623-176-46
5	37	88.1	9	15	US-10-623-176-2
6	37	88.1	9	15	US-10-791-217-2
7	37	88.1	9	17	US-10-861-335-1
8	36	85.7	9	15	US-10-623-176-1
9	36	85.7	9	16	US-10-791-217-1
10	34	81.0	9	15	US-10-623-176-48
11	33	78.6	9	15	US-10-623-176-23
12	33	78.6	9	15	US-10-623-176-41
13	33	78.6	9	15	US-10-623-176-45

14	29	69.0	9	15	US-10-623-176-5	Sequence 5, Appli
15	29	69.0	9	15	US-10-623-176-47	Sequence 47, Appl
16	28	66.7	9	15	US-10-623-176-40	Sequence 40, Appl
17	25	59.5	9	9	US-09-834-765-349	Sequence 349, App
18	25	59.5	9	15	US-10-623-176-39	Sequence 39, Appl
19	24	57.1	7	14	US-10-271-708-16	Sequence 16, Appl
20	24	57.1	9	15	US-10-623-176-4	Sequence 4, Appli
21	23	54.8	8	15	US-10-623-176-15	Sequence 15, Appl
22	23	54.8	9	9	US-09-834-765-132	Sequence 132, App
23	23	54.8	9	15	US-10-623-176-3	Sequence 3, Appli
24	22	52.4	9	15	US-10-623-176-14	Sequence 14, Appl
25	22	52.4	9	16	US-10-791-217-4	Sequence 4, Appli
26	21	50.0	7	10	US-09-990-186-208	Sequence 208, App
27	21	50.0	7	10	US-09-990-186-211	Sequence 211, App
28	21	50.0	7	10	US-09-990-186-216	Sequence 216, App
29	21	50.0	7	10	US-09-989-994-208	Sequence 208, App
30	21	50.0	7	10	US-09-989-994-211	Sequence 211, App
31	21	50.0	7	10	US-09-989-994-216	Sequence 216, App
32	21	50.0	7	14	US-10-122-675-6	Sequence 6, Appli
33	21	50.0	9	10	US-09-935-430-219	Sequence 219, App
34	21	50.0	9	10	US-09-935-430-565	Sequence 565, App
35	21	50.0	9	10	US-09-833-203-54	Sequence 54, Appl
36	21	50.0	9	14	US-10-119-536A-109	Sequence 109, App
37	21	50.0	9	14	US-10-277-292-219	Sequence 219, App
38	21	50.0	9	14	US-10-277-292-565	Sequence 565, App
39	21	50.0	9	15	US-10-280-340-219	Sequence 219, App
40	21	50.0	9	15	US-10-280-340-565	Sequence 565, App
41	21	50.0	9	15	US-10-057-475B-10608	Sequence 10608, A
42	21	50.0	9	15	US-10-057-475B-10813	Sequence 10813, A
43	21	50.0	9	15	US-10-154-884B-10608	Sequence 10608, A
44	21	50.0	9	15	US-10-154-884B-10813	Sequence 10813, A
45	21	50.0	9	15	US-10-394-980-421	Sequence 421, App

ALIGNMENTS

RESULT 1
US-10-623-176-10
; Sequence 10, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Els A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(9)
US-10-623-176-10

Query Match 100.0%; Score 42; DB 15; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.3e+06; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 VLRDDLEA 9
| | | | |
Db 1 VLRDDLEA 9

RESULT 2

US-10-791-217-5
; Sequence 5, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-4285US
; CURRENT APPLICATION NUMBER: US/10/791,217
; PRIOR FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: obtained from KIAA0223 partial complementary DNA
US-10-791-217-5

Query Match 100.0%; Score 42; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRDDLEA 9
| | | | |
Db 1 VLRDDLEA 9

RESULT 3

US-10-623-176-42
; Sequence 42, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(9)
US-10-623-176-42

Query Match 90.5%; Score 38; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLRDDLE 8
| | | | |
Db 2 VLRDDLE 9

RESULT 4

US-10-623-176-46
; Sequence 46, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; NAME/KEY: SITE
; LOCATION: (1)..(9)
US-10-623-176-46

Query Match 90.5%; Score 38; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LRDDLEA 9
| | | | |
Db 1 LRDDLEA 8

RESULT 5

US-10-623-176-2
; Sequence 2, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
US-10-623-176-2

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 07:32:15 ; Search time 23 Seconds
(without alignments)
29.211 Million cell updates/sec

Title: US-10-623-176-10

Perfect score: 42

Sequence: 1 VLRDDLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 99282

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgm2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgm2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgm2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgm2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgm2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgm2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	4	US-09-269-250E-18
2	37	88.1	9	4	US-09-269-250E-20
3	36	85.7	9	4	US-09-269-250E-29
4	27	64.3	9	3	US-09-217-609A-11
5	27	64.3	9	3	US-08-873-235B-11
6	23	54.8	8	4	US-09-269-250E-38
7	22	52.4	7	3	US-08-556-419-14
8	21	50.0	6	2	US-08-459-568-20
9	21	50.0	6	2	US-08-399-411-20
10	21	50.0	6	3	US-08-516-859A-20
11	21	50.0	6	3	US-09-586-472-20
12	21	50.0	6	4	US-09-528-706-20
13	21	50.0	8	2	US-08-459-568-75
14	21	50.0	8	2	US-08-399-411-75
15	21	50.0	8	2	US-08-539-432-1
16	21	50.0	8	3	US-08-516-859A-75
17	21	50.0	8	3	US-09-586-472-75
18	21	50.0	8	4	US-09-528-706-75
19	21	50.0	8	4	US-08-453-485E-13
20	21	50.0	9	2	US-08-459-568-72
21	21	50.0	9	2	US-08-459-568-78
22	21	50.0	9	2	US-08-399-411-72
23	21	50.0	9	2	US-08-399-411-78
24	21	50.0	9	3	US-08-516-859A-72
25	21	50.0	9	3	US-08-516-859A-78
26	21	50.0	9	3	US-09-586-472-72
27	21	50.0	9	3	US-09-586-472-78

```

28 21 50.0 9 4 US-09-528-706-72 Sequence 72, Appl
29 21 50.0 9 4 US-09-528-706-78 Sequence 78, Appl
30 21 50.0 9 4 US-09-744-549-15 Sequence 15, Appl
31 20 47.6 6 3 US-09-025-819-11 Sequence 11, Appl
32 20 47.6 6 4 US-09-808-126-11 Sequence 11, Appl
33 20 47.6 6 4 US-09-803-951-11 Sequence 11, Appl
34 20 47.6 8 3 US-08-747-599A-47 Sequence 47, Appl
35 20 47.6 8 3 US-08-747-599A-50 Sequence 50, Appl
36 20 47.6 8 3 US-08-747-599A-51 Sequence 51, Appl
37 20 47.6 9 4 US-08-634-332A-50 Sequence 50, Appl
38 19 45.2 7 3 US-09-187-859-1186 Sequence 1186, Ap
39 19 45.2 7 4 US-09-839-542B-1186 Sequence 1186, Ap
40 19 45.2 8 3 US-09-187-859-1189 Sequence 1189, Ap
41 19 45.2 8 3 US-09-183-266A-38 Sequence 38, Appl
42 19 45.2 8 4 US-09-839-542B-1189 Sequence 1189, Ap
43 19 45.2 9 1 US-08-467-083-20 Sequence 20, Appl
44 19 45.2 9 1 US-08-414-417B-20 Sequence 20, Appl
45 19 45.2 9 2 US-08-486-348A-20 Sequence 20, Appl

```

ALIGNMENTS

RESULT 1

```

US-09-269-250E-18
; Sequence 18, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Exon Fragments
US-09-269-250E-18

```

Query Match 100.0%; Score 42; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 VLRDDLEA 9
Db 1 VLRDDLEA 9

```

RESULT 2

```

US-09-269-250E-20
; Sequence 20, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Exon Fragments
US-09-269-250E-20

```

Query Match 88.1%; Score 37; DB 4; Length 9;

Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLKDDLLLEA 9
Db 1 VLKDDLLLEA 9

RESULT 3
US-09-269-250E-29
; Sequence 29, Application US/09269250E
; Patent No. 6830883
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elsa
; TITLE OF INVENTION: METHOD FOR TYPING OF MINOR HISTOCOMPATIBILITY ANTIGEN HA-1
; FILE REFERENCE: 58994
; CURRENT APPLICATION NUMBER: US/09/269,250E
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3)-(3)
; OTHER INFORMATION: Xaa represents a histidine (H) or an arginine (R) residue

US-09-269-250E-29

Query Match 85.7%; Score 36; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLKDDLLLEA 9
Db 1 VLKDDLLLEA 9

RESULT 4
US-09-217-609A-11
; Sequence 11, Application US/09217609A
; Patent No. 6071733
; GENERAL INFORMATION:
; APPLICANT: MURAMATSU, Masayoshi
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OGURA, Kyoza
; APPLICANT: Koyama, Tanetoshi
; APPLICANT: SHIMIZU, Naoto
; APPLICANT: CHO, Yenwin
; TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, NW - Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: WordPerfect 6.1 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/217,609A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/873,235
; FILING DATE: 11-Jun-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: TOFFENETTI, Judith L.

; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 10235/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-1776
; TELEFAX: 202-429-0796
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-217-609A-11

Query Match 64.3%; Score 27; DB 3; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LRDDLLLE 8
Db 3 IRDDILD 9

RESULT 5
US-08-873-235B-11
; Sequence 11, Application US/08873235B
; Patent No. 6174715
; GENERAL INFORMATION:
; APPLICANT: MURAMATSU, Masayoshi
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OGURA, Kyoza
; APPLICANT: Koyama, Tanetoshi
; APPLICANT: SHIMIZU, Naoto
; APPLICANT: CHO, Yenwin
; TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, NW - Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: WordPerfect 6.1 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,235B
; FILING DATE: 11-Jun-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 154441/1996
; FILING DATE: 14-Jun-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: TOFFENETTI, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 10235/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-1776
; TELEFAX: 202-429-0796
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-873-235B-11

Query Match 64.3%; Score 27; DB 3; Length 9;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;